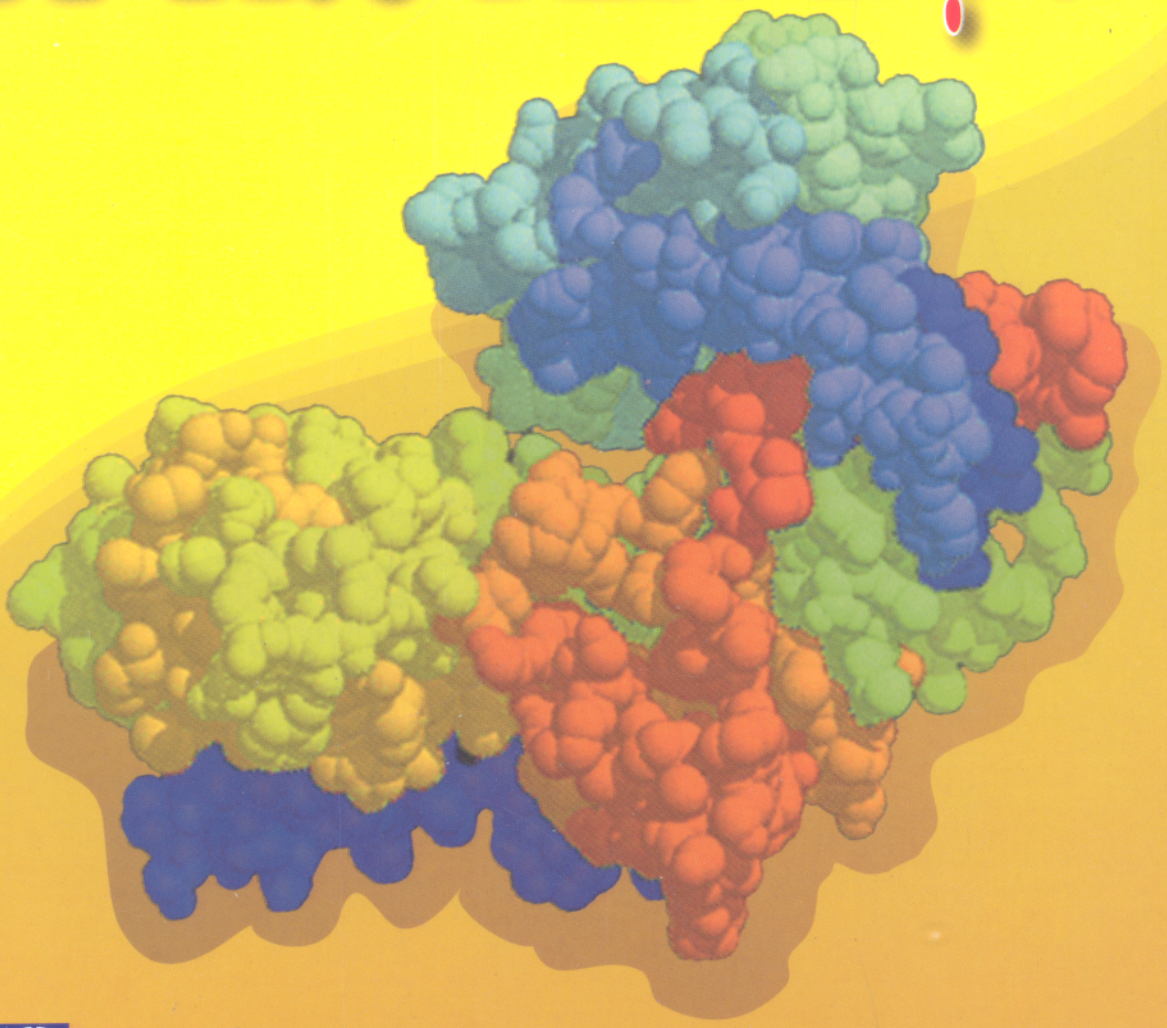
ĐỖ LÊ THĂNG – ĐINH ĐOÀN LONG

Chú giải
DI TRUYỀN HỌC



NHÀ XUẤT BẢN GIÁO DỤC

ĐỖ LÊ THĂNG – ĐINH ĐOÀN LONG

CHÚ GIẢI
DI TRUYỀN HỌC

NHÀ XUẤT BẢN GIÁO DỤC

Bản quyền thuộc HEVOBCO – Nhà xuất bản Giáo dục

17 – 2007/CXB/91 – 2217/GD

Mã số: 7K701M7 – D

Lời nói đầu

Sách "*Chú giải Di truyền học*" được biên soạn nhằm giúp người đọc dễ dàng tra cứu các vấn đề của Di truyền học theo các chủ đề lớn và theo trình tự logic các sự kiện của vấn đề. Bố cục của cuốn sách còn giúp người đọc hình dung các quá trình sinh học từ mức phân tử đến quần thể.

Cuốn sách gồm các phần sau:

Phần I. Di truyền học phân tử

Phần II. Hệ gen

Phần III. Các cơ chế di truyền

Phần IV. Di truyền học quần thể và tiến hóa

Phần V. Công nghệ ADN tái tổ hợp

Phần VI. Các ứng dụng của di truyền học.

Ngoài những kiến thức cơ bản, nhiều mục của cuốn sách đã được cập nhật những thông tin mới nhất, như mục *"Dự án hệ gen người", Tiến hoá phân tử"*... Phần "*Các ứng dụng của di truyền học*" đã tổng kết các lĩnh vực ứng dụng chính của di truyền học trong nghiên cứu cơ bản, tiến hóa, y học và đời sống xã hội.

Chúng tôi tin tưởng rằng, cuốn sách sẽ là tài liệu tham khảo tốt cho những người làm công tác giảng dạy di truyền học ở các Trường đại học, cao đẳng và nghiên cứu khoa học. Sách còn là tài liệu tham khảo hữu ích cho các thày cô giảng dạy môn Sinh học và học sinh các Trường phổ thông trung học.

Trong quá trình biên dịch, chắc chắn còn có thiếu sót. Chúng tôi mong nhận được các ý kiến đóng góp của bạn đọc để lần tái bản sau sách được hoàn thiện hơn. Mọi ý kiến xây dựng xin được gửi về Công ty Cổ phần Sách Đại học – Dạy nghề, 25 Hà Thuyên, Hà Nội. Điện thoại (04)8266360.

Xin chân thành cảm ơn các bạn đọc.

CÁC TÁC GIẢ

MỤC LỤC

Lời nói đầu 3

Phần I
DI TRUYỀN HỌC PHÂN TỬ

I.A ADN (acid deoxyribonucleic) 7
I.B Gen 11
I.C Mã di truyền 14
I.D Phiên mã 16
I.E ARN vận chuyển (tARN) 21
I.F ARN ribosom (rARN) 24
I.G ARN thông tin (mARN) 27
I.H Dịch mã 31
I.I Sao chép ADN 38
I.K Điều hòa biểu hiện của gen ở sinh vật nhân sơ 44
I.L Điều hòa biểu hiện gen ở sinh vật nhân chuẩn 50

Phần II
NHIỄM SẮC THỂ VÀ HỆ GEN

II.A Nhiễm sắc thể 56
II.B Sự phân chia tế bào (nguyên phân) 63
II.C Hệ gen sinh vật nhân sơ 65
II.D Hệ gen người 69
II.E Đột biến ADN 73
II.F Các tác nhân đột biến và sửa chữa ADN 77
II.G Tái tổ hợp 82
II.H Phage 88
II.I Virus lây nhiễm sinh vật nhân chuẩn 92

Phần III
CÁC CƠ CHẾ DI TRUYỀN

III. A Di truyền học Mendel 99
III.B Di truyền học Mendel mở rộng 105
III.C Giảm phân và quá trình phát sinh giao tử 111
III.D Liên kết gen 116
III.E Sự truyền gen ở vi khuẩn 122

III.F Sự di truyền ngoài nhân 128
III.G Sự di truyền các tính trạng số lượng 131
III.H Sự xác định giới tính 139
III.I Sự di truyền liên kết giới tính 143
III.K Lai tế bào soma 145
III.L Nội phối 150
III.M Xác suất 154
III.N Phương pháp χ^2 và kiểm tra giả thiết 158

Phần IV
DI TRUYỀN HỌC QUẦN THỂ VÀ TIẾN HÓA

IV.A Giới thiệu 164
IV.B Tiến hóa nhờ chọn lọc tự nhiên 166
IV.C Trạng thái cân bằng Hardy-Weinberg 171
IV.D Đa dạng di truyền 176
IV.E Thuyết Đarwin mới: chọn lọc tác động đến các alen 180
IV.F Sự biến đổi nhiễm sắc thể trong quá trình tiến hóa 183
IV.G Loài và sự hình thành loài 192
IV.H Đa bội thể 197
IV.I Quá trình tiến hóa 201

Phần V
CÔNG NGHỆ ADN TÁI TỔ HỢP

V.A Lai acid nucleic 208
V.B Tách dòng ADN 212
V.C PCR - Phản ứng chuỗi trùng hợp 221
V.D Giải mã trình tự ADN 225

Phần VI
CÁC ỨNG DỤNG CỦA DI TRUYỀN HỌC

VI. A Các bệnh di truyền 231
VI.B Di truyền học và ung thư 239
VI.C Phép thử độc tính gen 244
VI.D Liệu pháp gen 247
VI.E Dự án hệ gen người 251
VI.F Di truyền học và khoa học hình sự 257
VI.G Kỹ thuật di truyền và công nghệ sinh học 264
TÀI LIỆU THAM KHẢO 274

Phần I
DI TRUYỀN HỌC PHÂN TỬ

Phần I gồm các mục:

I.A ADN
- Các nucleotid
- Chuỗi polynucleotid ADN
- Chuỗi xoắn kép
- Nguyên tắc ghép đôi các base bổ trợ
- Cấu trúc ARN

I.B Gen
- Cấu trúc của gen
- Họ gen
- Sự biểu hiện của gen
- Gen khởi động
- Intron và exon
- Gen giả

I.C Mã di truyền
- Sự biểu hiện gen
- Mã di truyền
- Khung đọc
- Tính vạn năng của mã di truyền

I.D Phiên mã
- Phiên mã
- Phiên mã ở sinh vật nhân sơ
- Sự khởi đầu phiên mã
- Kéo dài phiên mã
- Kết thúc phiên mã
- Phiên mã ở sinh vật nhân chuẩn
- ARN polymerase I
- ARN polymerase II
- ARN polymerase III

I.E ARN vận chuyển (tARN)
- Vai trò của tARN trong quá trình dịch mã
- Cấu trúc tARN
- Tổng hợp và hoàn thiện
- Sửa đổi các nucleotid

I.F ARN ribosom (rARN)
- Ribosom
- Phiên mã và hoàn thiện các rARN ở sinh vật nhân sơ
- Phiên mã và hoàn thiện các rARN ở sinh vật nhân chuẩn

I.G ARN thông tin (mARN)
- Tổng hợp và hoàn thiện mARN
- Cắt bỏ các intron
- Lắp mũ
- Gắn đuôi poly(A)
- Tính bền vững của mARN
- Các phân tử mARN được hoàn thiện theo các cách khác nhau

I.H Dịch mã
- Vai trò của tARN
- Nhận biết codon
- Dịch mã
- Khởi đầu dịch mã
- Kéo dài dịch mã
- Kết thúc dịch mã
- Sửa đổi sau dịch mã

I.I Sao chép ADN
- Sao chép ADN
- Chạc sao chép
- Sao chép ADN ở vi khuẩn
- Sao chép ADN ở sinh vật nhân chuẩn

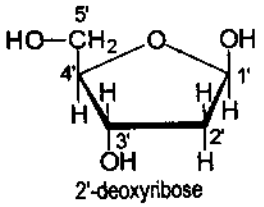
I.K Điều hòa biểu hiện của gen ở sinh vật nhân sơ
- Điều hòa sự biểu hiện của gen
- Tổ chức của các gen vi khuẩn
- Operon lac
- Ức chế bởi chất dị hóa
- Operon trp
- Phiên mã không hoàn chỉnh
- Điều hòa bằng các yếu tố σ thay thế

I.L Điều hòa biểu hiện gen ở sinh vật nhân chuẩn
- Sự biểu hiện của gen nhân chuẩn
- Điều hòa phiên mã
- Các yếu tố phiên mã
- Điều hòa hoạt động gen nhờ hormon và các hợp chất cytokin

I.A ADN (ACID DEOXYRIBONUCLEIC)

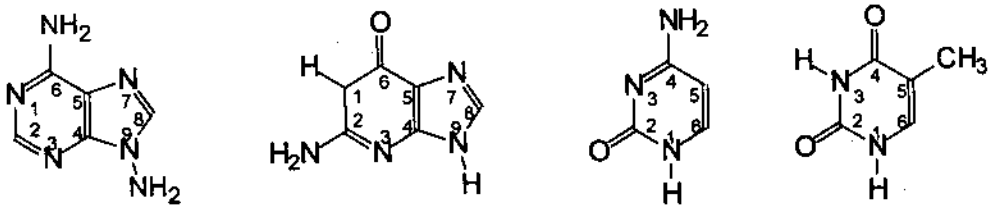
Các nucleotid

Khả năng mang thông tin di truyền của phân tử ADN liên quan chặt chẽ với cấu trúc phân tử của nó. ADN là một đại phân tử dạng polyme gồm nhiều tiểu phần (monome) là các **nucleotid**. Vì vậy, các phân tử ADN còn được gọi là các polynuleotid. Mỗi một nucleotid gồm có ba thành phần: một gốc đường, một base có cấu trúc vòng chứa nitơ (gọi là **base nitơ**) và một nhóm phosphat. Thành phần đường trong phân tử ADN là đường năm carbon (pentose, còn được gọi là 2'- deoxyribose), trong đó nhóm -OH ở vị trí carbon số 2' của đường ribose được thay thế bằng nguyên tử hydro (*Hình 1*). Các nguyên tử carbon trong đường được đánh số thứ tự từ 1 đến 5. Các số này được đánh dấu **phảy** (và thường được gọi là C-1' đến C-5') để phân biệt với các vị trí carbon trong thành phần base nitơ. Việc đánh số là cần thiết để giúp chỉ ra vị trí gắn của các thành phần khác trên thành phần đường của nucleotid.



2'-deoxyribose

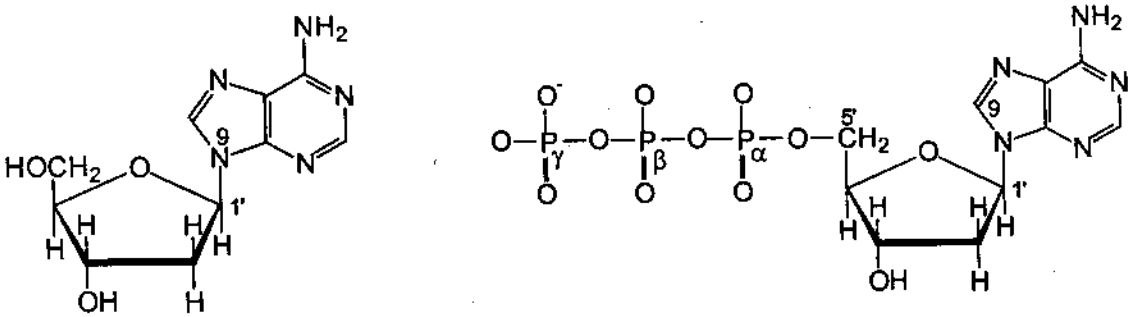
Hình 1.1. **Cấu trúc của 2'-deoxyribose**



Adenin (A) Guanin (G) Cytosin (C) Thymin (T)

Hình 1.2. **Các loại base nitơ của ADN**

Các nucleotid chứa một trong bốn loại base nitơ: **adenin, guanin, thymin** và **cytosin** (*Hình 1.2*). Đây là thành phần có cấu trúc hóa học phức tạp chứa các cấu trúc vòng carbon và nitơ. Adenin và guanin chứa hai vòng carbon-nitơ và được gọi là các **purin**. Cytosin và thymin chứa một cấu trúc vòng và được gọi là các **pyrimidin**. Các base gắn vào phân tử đường thông qua liên kết giữa C-1' của đường với nitơ ở vị trí số 9 của purin hoặc vị trí số 1 của pyrimidin. Phần cấu trúc gồm đường và base nitơ được gọi là **nucleozid** (*Hình 1.3a*).

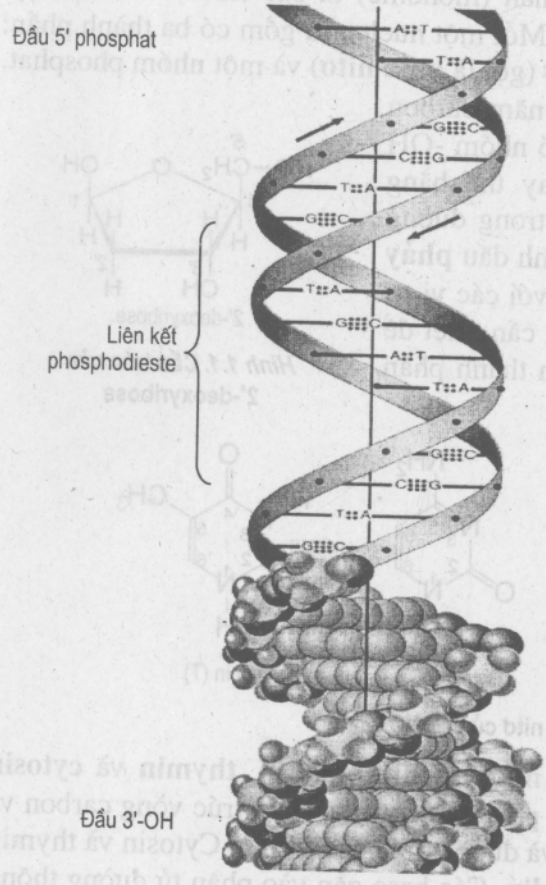


a) 2'-deoxyadenosin (nucleozid) b) 2'-deoxyadenosin 5'-triphosphat (dATP) (nucleotid)

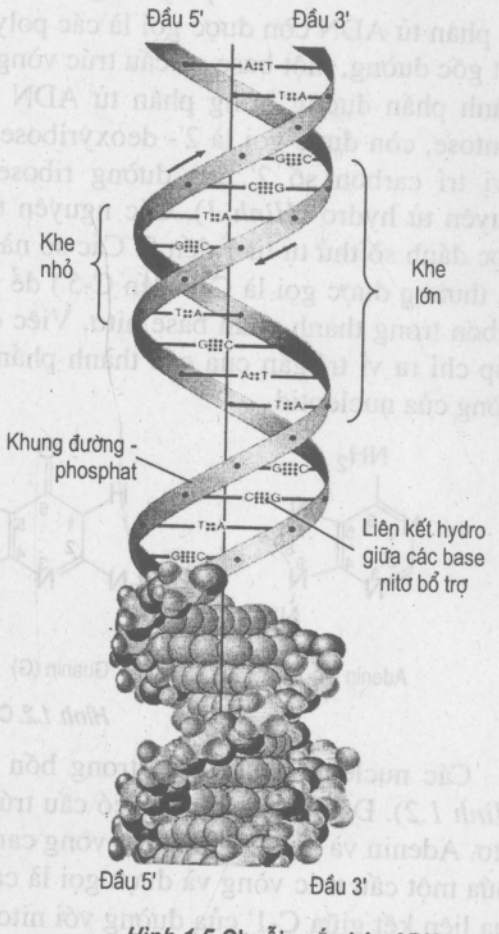
Hình 1.3. **Cấu trúc của nucleozid (a) và nucleotid (b)**

Các nucleotid chứa nhóm phosphat (PO_4) gắn vào vị trí C-5' của đường (*Hình 1.3b*). Như vậy, một nucleozid sẽ được gọi là nucleotid khi được gắn thêm nhóm phosphat. Phần gắn thêm có thể mang một, hai hoặc ba nhóm phosphat, với thứ tự các nhóm phosphat được gọi là

α, β và γ kể từ vị trí liên kết với đường. Trong tế bào, các phân tử nucleotid có thể tồn tại như những phân tử độc lập (các nucleotid triphosphat có một vai trò quan trọng là phân tử mang năng lượng cao năng cần thiết cho nhiều hoạt động khác nhau của tế bào) hoặc là tiểu phần cấu thành nên các đại phân tử (polyme) sinh học là các acid nucleic (ADN và ARN).



Hình 1.4. **Liên kết phosphodieste gắn kết các nucleotid trong chuỗi polynucleotid**

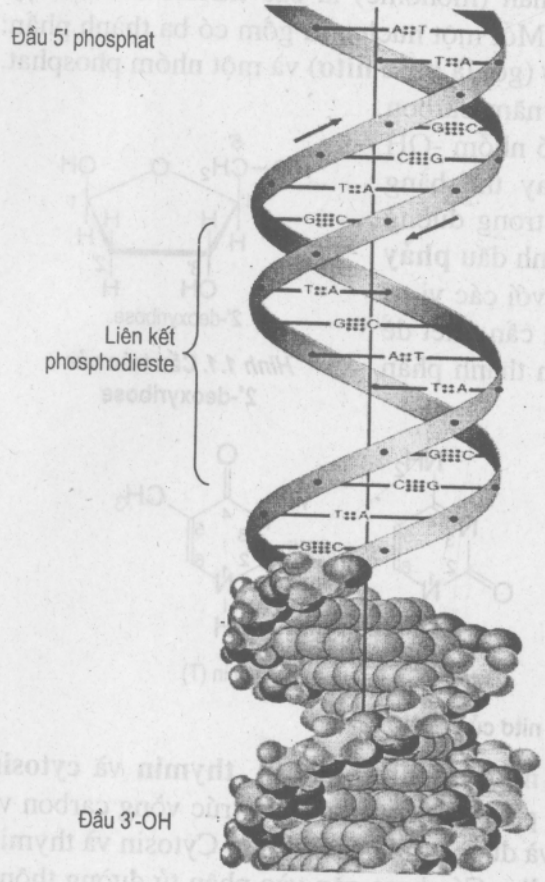


Hình 1.5. **Chuỗi xoắn kép ADN**

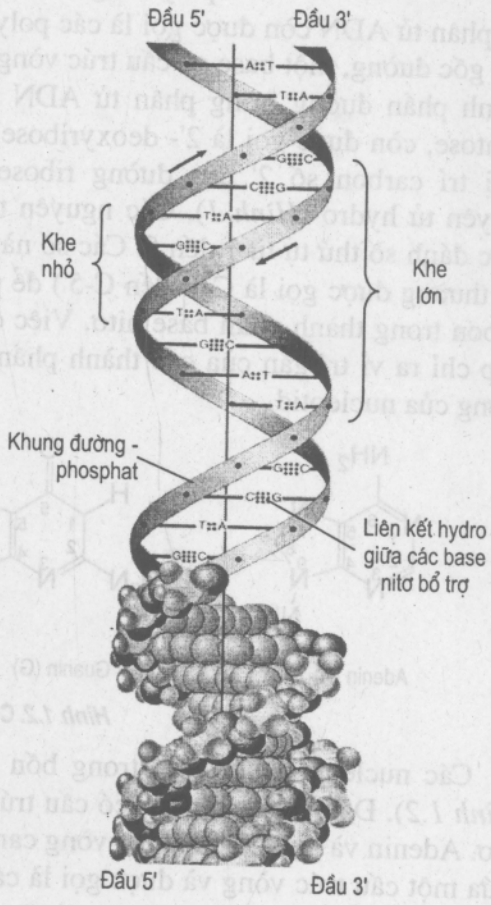
Chuỗi polynucleotid

Các nucleotid triphosphat liên kết với nhau tạo ra chuỗi polynucleotid. Có bốn loại nucleotid triphosphat được sử dụng để tổng hợp nên chuỗi polynucleotid là: 2'-deoxyadenosine 5'-triphosphat (viết tắt là dATP hoặc A), 2'-deoxythymidin 5'-triphosphat (dTTP hoặc T), 2'-deoxycytosin 5'-triphosphat (dCTP hoặc C) và 2'-deoxyguanosine 5'-triphosphat (dGTP hoặc G). Các nhóm phosphat β và γ bị mất đi trong quá trình polyme hoá và các nucleotid kết hợp với nhau nhờ nhóm phosphat (α) còn lại. Phosphat ở đầu 5' của nucleotid này hình thành mối liên kết với carbon 3' của nucleotid kế tiếp thông qua việc loại bỏ nhóm -OH ở vị trí carbon 3' trong quá trình tổng hợp chuỗi ADN. Mối liên kết này được gọi là liên kết **phosphodieste 3'-5'** (C-O-P) (*Hình 1.4*). Chuỗi polynucleotid có 5' triphosphat tự do ở một đầu (được gọi là đầu 5') và nhóm 3' hydroxyl ở đầu kia (đầu 3'). Điều đó làm cho chuỗi ADN ở mỗi mạch có tính phân cực, và thường được mô tả là chạy theo chiều 5' → 3' hoặc 3'→ 5'.

α, β và γ kể từ vị trí liên kết với đường. Trong tế bào, các phân tử nucleotid có thể tồn tại như những phân tử độc lập (các nucleotid triphosphat có một vai trò quan trọng là phân tử mang năng lượng cao năng cần thiết cho nhiều hoạt động khác nhau của tế bào) hoặc là tiểu phần cấu thành nên các đại phân tử (polyme) sinh học là các acid nucleic (ADN và ARN).



Hình 1.4. **Liên kết phosphodieste gắn kết các nucleotid trong chuỗi polynucleotid**



Hình 1.5. **Chuỗi xoắn kép ADN**

Chuỗi polynucleotid

Các nucleotid triphosphat liên kết với nhau tạo ra chuỗi polynucleotid. Có bốn loại nucleotid triphosphat được sử dụng để tổng hợp nên chuỗi polynucleotid là: 2'-deoxyadenosine 5'-triphosphat (viết tắt là dATP hoặc A), 2'-deoxythymidin 5'-triphosphat (dTTP hoặc T), 2'-deoxycytosin 5'-triphosphat (dCTP hoặc C) và 2'-deoxyguanosine 5'-triphosphat (dGTP hoặc G). Các nhóm phosphat β và γ bị mất đi trong quá trình polyme hoá và các nucleotid kết hợp với nhau nhờ nhóm phosphat (α) còn lại. Phosphat ở đầu 5' của nucleotid này hình thành mối liên kết với carbon 3' của nucleotid kế tiếp thông qua việc loại bỏ nhóm -OH ở vị trí carbon 3' trong quá trình tổng hợp chuỗi ADN. Mối liên kết này được gọi là liên kết **phosphodieste 3'-5'** (C-O-P) (*Hình 1.4*). Chuỗi polynucleotid có 5' triphosphat tự do ở một đầu (được gọi là đầu 5') và nhóm 3' hydroxyl ở đầu kia (đầu 3'). Điều đó làm cho chuỗi ADN ở mỗi mạch có tính phân cực, và thường được mô tả là chạy theo chiều 5' → 3' hoặc 3'→ 5'.

Trình tự các base trên ADN mã hoá thông tin di truyền. Trình tự mã hóa luôn được đọc theo chiều 5'→ 3'(do enzym ADN polymerase chỉ sao chép phân tử ADN theo hướng này). Các chuỗi polynucleotid có thể rất dài với số lượng các nucleotid không hạn chế và trật tự các nucleotid không giới hạn. Số kiểu trình tự sắp xếp tối đa của một chuỗi polynucleotid có thể có là 4^n với n là số nucleotid. Ví dụ, một chuỗi polynucleotid gồm 6 base có thể có $4^6 = 4096$ kiểu trình tự sắp xếp khác nhau.

Chuỗi xoắn kép

Phân tử ADN có cấu trúc không gian rất đặc trưng và được gọi là chuỗi xoắn kép (*Hình 1.5*). Cấu trúc ADN được Watson và Crick phát hiện năm 1953 tại Cambridge (Anh) nhờ sử dụng các ảnh nhiễu xạ tia X trước đó được chụp bởi Franklin và Wilkins. ADN gồm hai chuỗi polynucleotid quấn quanh nhau tạo nên chuỗi xoắn kép. Phần đường-phosphat hình thành nên bộ khung phía ngoài của chuỗi xoắn. Các base là những phân tử mỏng, nằm ở phía trong, hướng vào tâm bộ khung này và xếp chồng lên nhau như một chồng đĩa.

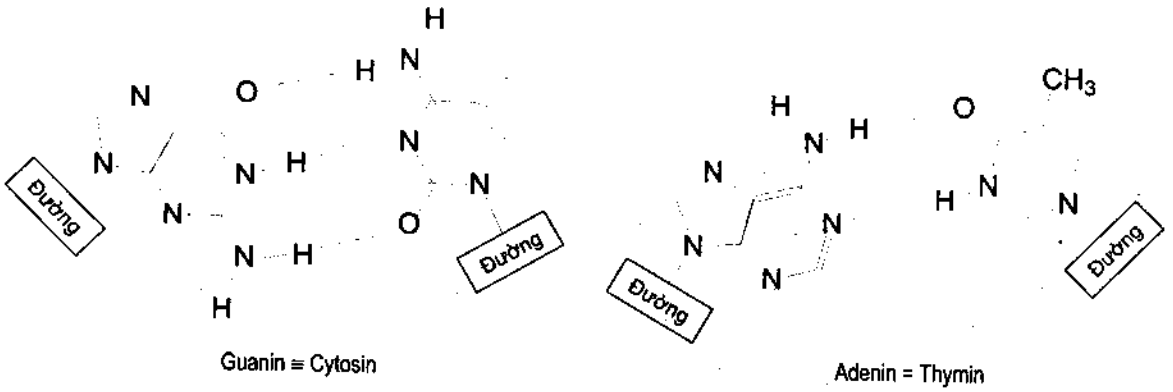
Các ảnh nhiễu xạ tia X cho thấy hình ảnh các băng lặp đi lặp lại phản ánh tính cân đối của cấu trúc ADN. Chuỗi xoắn kép xoắn thành một vòng hoàn chỉnh sau mỗi 10 cặp base. Độ dài một bước xoắn là 34Å nên khoảng không giữa các base là 3,4Å. Đường kính chuỗi xoắn là 20Å. Chuỗi xoắn kép được gọi là chuỗi **đối song song** (hai mạch chạy song song ngược chiều). Một mạch chạy theo chiều 5'→ 3', còn mạch kia theo chiều 3'→ 5'. Chỉ những chuỗi polynucleotid đối song song mới tạo nên chuỗi xoắn bền vững. Chuỗi xoắn kép không hoàn toàn đều đặn và nếu nhìn từ ngoài có thể thấy **rãnh lớn** (khe lớn) và **rãnh nhỏ** (khe nhỏ). Những rãnh này rất quan trọng để tương tác với các protein điều hòa và protein cấu trúc trong quá trình sao chép và biểu hiện của thông tin di truyền. Chuỗi xoắn kép ADN phổ biến là xoắn về bên phải, nghĩa là, nếu tưởng tượng chuỗi xoắn kép giống một cầu thang xoắn mà khi bạn leo lên cầu thang đó thì khung đường-phosphat ở phía bên tay phải của bạn. Kiểu cấu hình không gian này của phân tử ADN sau khi được tinh thể hóa được gọi là **dạng B**. Trong những điều kiện môi trường nhất định, phân tử ADN có thể có một số dạng cấu hình không gian khác. Chẳng hạn, dạng A có cấu trúc chặt hơn. Một số dạng cấu hình khác là C, D, E và Z (còn gọi là dạng S) có cấu trúc xoắn trái. Đến nay đã có 21 dạng cấu hình không gian khác nhau của ADN được mô tả (trong bảng chữ cái, hiện nay chỉ còn các chữ cái F, Q, U, V và Y chưa được dùng để chỉ các dạng cấu hình không gian của ADN).

Nguyên tắc ghép đôi các base bổ trợ

Các base của hai chuỗi polynucleotid liên kết với nhau. Khoảng không gian giữa hai chuỗi polynucleotid chỉ đủ cho một vòng purin kết hợp với một vòng pyrimidin. Như vậy, thymin (T) luôn liên kết với adenin (A); và guanin (G) luôn liên kết với cytosin (C). Các liên kết hydro hình thành giữa các base và làm ổn định mối tương tác đó. Hai liên kết hydro được hình thành giữa A và T, và ba liên kết giữa G và C. Như vậy, sự liên kết giữa G-C bền vững hơn A-T. Tuy vậy, để các liên kết hydro có thể nằm trên một "mặt phẳng", các tiểu phần phân tử nucleotid (giữa A và T, và giữa G và C) phải quay ngược nhau 180°. Cách các base ghép cặp giữa hai sợi ADN được gọi là nguyên tắc ghép đôi của các base bổ trợ (được gọi tắt là nguyên tắc bổ trợ, hay nguyên tắc Chargaff) và có ý nghĩa đặc biệt quan trọng

(*Hình 1.6*). Vì G luôn liên kết với C và A liên kết với T nên trình tự các nucleotid của hai mạch trên phân tử ADN sợi kép liên quan chặt chẽ với nhau và được gọi là bổ trợ cho nhau, vì trình tự của mạch này xác định trình tự của mạch kia. Điều đó có nghĩa là có thể dùng một mạch để sao chép mạch kia. Đây chính là cơ chế sống còn để duy trì thông tin di truyền và truyền nó cho các tế bào khác trong quá trình phân bào. Nguyên tắc ghép đôi của các base bổ trợ còn quan trọng trong quá trình biểu hiện thông tin di truyền và là nguyên lý cơ bản để phiên mã các trình tự deoxyribonucleotid của phân tử ADN thành trình tự ribonucleotid của mARN và dịch mã thành trình tự các amino acid (axit amin) trong phân tử protein tương ứng.

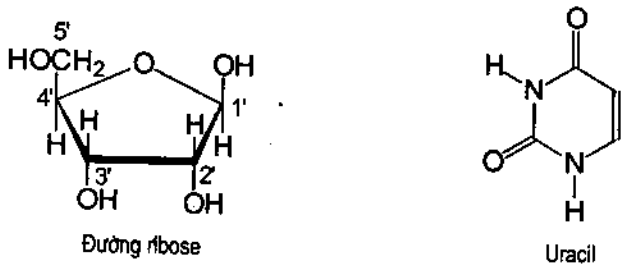
Các liên kết hydro có thể bị phá vỡ bởi nhiệt hoặc một số hoá chất. Kết quả là chuỗi xoắn kép ADN tách ra thành hai mạch đơn và hiện tượng đó được gọi là sự biến tính ADN. Trong tế bào, các enzym có thể tách hai mạch đơn trong quá trình sao chép (tái bản) ADN và biểu hiện thông tin di truyền.



Hình 1.6. **Các cặp base nitơ liên kết bổ trợ (liên kết hydro được minh họa bằng vạch ngắt quãng)**

Cấu trúc ARN

ARN có cấu trúc tương tự ADN, song có một số khác biệt quan trọng. Trong ARN, đường ribose thay thế cho 2'-deoxyribose và base thymin được thay bằng uracil là base cũng thường bắt cặp với adenin (*Hình 1.7*). Hơn nữa, phân tử ARN thường tồn tại ở dạng một mạch polynucleotid đơn lẻ và không hình thành chuỗi xoắn kép. Tuy nhiên, có thể có sự ghép đôi giữa các phần bổ trợ giữa các sợi ARN và giữa ARN và ADN, tạo nên những vùng có cấu trúc sợi kép ngắn.



Hình 1.7. **Cấu trúc của đường ribose và uracil (U)**

I.B GEN

Khái niệm gen

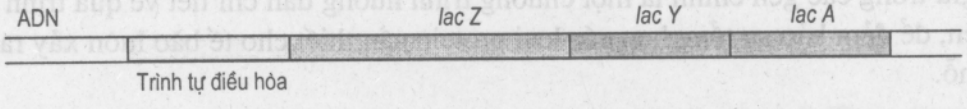
Các thông tin di truyền mà sinh vật cần cho quá trình sinh trưởng, phát triển và sinh sản nằm trong phân tử ADN của nó. Những thông tin này nằm trong trình tự nucleotid của ADN và được tổ chức thành các gen. Mỗi gen thường chứa thông tin để tổng hợp một chuỗi polypeptid hoặc một phân tử ARN có chức năng riêng biệt. Xét về mặt cấu trúc, mỗi gen là một đoạn ADN riêng biệt mang trình tự base thường mã hoá cho trình tự amino acid của một chuỗi polypeptid. Các gen rất khác nhau về kích thước, có thể từ dưới 100 cặp đến vài triệu cặp base. Ở sinh vật bậc cao, các gen hợp thành các phân tử ADN rất dài nằm trong các cấu trúc được gọi là **nhiễm sắc thể**. Ở người có khoảng 30.000 - 40.000 gen phân bố trên 23 cặp nhiễm sắc thể. Trên nhiễm sắc thể, các gen thường nằm phân tán và cách biệt nhau bởi các đoạn trình tự không mã hóa. Các đoạn trình tự này được gọi là các đoạn **ADN liên gen**. ADN liên gen rất dài, như ở người các gen chỉ chiếm dưới 30% toàn bộ hệ gen. Xét ở mỗi gen, chỉ một mạch của chuỗi xoắn kép là mang thông tin và được gọi là **mạch khuôn** dùng để tạo ra phân tử ARN mang trình tự bổ trợ để điều khiển quá trình tổng hợp chuỗi polypeptid. Mạch kia được gọi là **mạch không làm khuôn**. Cả hai mạch trên phân tử ADN đều có thể được dùng làm mạch để mã hoá cho các gen khác nhau. Ngoài ra, người ta còn dùng một số thuật ngữ khác để chỉ mạch khuôn và mạch không làm khuôn, như **mạch mang nghĩa/mạch đối nghĩa**, **mạch mã hoá/mạch không mã hoá**. Cần chú ý là, *mạch đối nghĩa* và *mạch không mã hóa* chính là *mạch khuôn* để tổng hợp phân tử ARN.

Khả năng lưu giữ thông tin di truyền của ADN là rất lớn. Với một phân tử ADN có *n* base sẽ có 4^n các tổ hợp trình tự base khác nhau. Thậm chí với một phân tử ADN rất ngắn đã có rất nhiều kiểu trình tự khác nhau. Trong thực tế, chỉ một số lượng hạn chế các trình tự mang thông tin có ích (thông tin mã hóa các phân tử ARN hoặc protein có chức năng sinh học).

Tổ chức của gen

Hầu hết các gen phân bố ngẫu nhiên trên nhiễm sắc thể, tuy nhiên có một số gen được tổ chức thành nhóm, hoặc cụm. Có hai kiểu cụm gen, đó là các **operon** và các **họ gen**.

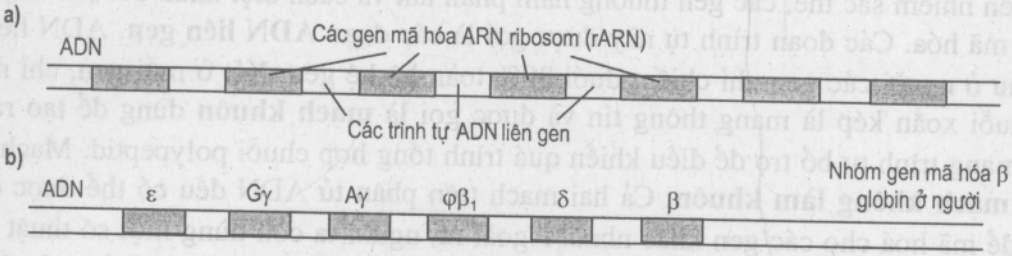
Operon là các cụm gen ở vi khuẩn. Chúng thường chứa các gen được điều hoà hoạt động đồng thời và mã hoá cho các protein có chức năng liên quan với nhau. Ví dụ như operon *lac* ở *E. coli* chứa ba gen mã hoá cho các enzym mà vi khuẩn cần để thủy phân lactose. Khi có lactose làm nguồn năng lượng (và vắng mặt glucose) thì vi khuẩn cần ba enzym do operon *lac* mã hoá. Sự cụm lại của các gen trong operon (*Hình 1.8*) cho phép các gen đó được điều khiển biểu hiện đồng thời để sinh vật có thể sử dụng nguồn năng lượng một cách hiệu quả.



Hình 1.8. Operon *Lac*. Ba gen (lac Z, *lac* Y và *lac* A) xếp liền kề nhau và được điều khiển chung

Ở các sinh vật bậc cao không có các operon, các cụm gen được gọi là các **họ gen**. Không giống như các operon, các gen trong một họ gen rất giống nhau, nhưng không được điều khiển

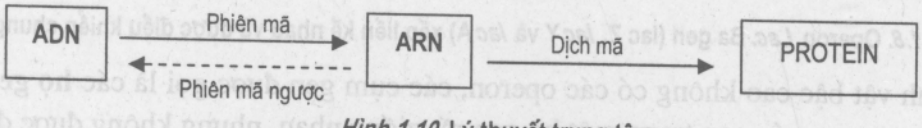
biểu hiện đồng thời. Sự cụm lại của các gen trong họ gen có lẽ phản ánh nhu cầu cần có nhiều bản sao của những gen nhất định trong quá trình tiến hoá, và thu được do sự lặp đoạn của gen. Một số họ gen tồn tại thành nhiều cụm riêng biệt trên nhiều nhiễm sắc thể khác nhau. Hiện tượng đó có lẽ là do sự tái cấu trúc ADN trong quá trình tiến hoá đã phá vỡ các cụm gen. Các họ gen có thể có cấu trúc đơn giản hoặc phức tạp. Ở các **họ gen đơn giản**, các gen giống hệt nhau. Ví dụ như họ gen mã hóa ARN ribosom 5S (rARN 5S). Ở mỗi tế bào người, có khoảng 2000 cụm gen của gen này, phản ánh tế bào cần số lượng lớn sản phẩm của gen này (*Hình 1.9a*). Trong khi đó, các **họ gen phức tạp** chứa các gen tương tự nhưng không giống hệt nhau. Ví dụ như họ gen globin ở người mã hóa cho cho các chuỗi polypeptid tương ứng với các loại globin α, β, γ, ε, và ζ (*Hình 1.9b*) chỉ khác nhau vài amino acid. Các chuỗi polypeptid globin tương tác với nhau thành một phức hệ, và kết hợp với các phân tử hem để tạo ra hemoglobin (một loại protein vận chuyển oxy trong máu).



Hình 1.9. **Một họ gen đơn giản (a), và một họ gen phức tạp (b)**

Sự biểu hiện của gen

Trình tự các base trong ADN chính là thông tin di truyền. Sự biểu hiện của gen là quá trình tế bào sử dụng các thông tin đó. Thông tin di truyền được sử dụng theo một nguyên tắc gọi là **lý thuyết trung tâm** do Crick đưa ra và được phát biểu rằng "thông tin di truyền được truyền từ ADN đến ARN, rồi đến protein" (*Hình 1.10*). Trong quá trình biểu hiện của gen, phân tử ADN phiên bản thông tin của mình bằng cách trực tiếp tổng hợp nên một phân tử ARN có trình tự các nucleotid bổ trợ. Quá trình này được gọi là **sự phiên mã**. Sau đó ARN điều khiển việc tổng hợp chuỗi polypeptid có trình tự các amino acid do trình tự base (nucleotid) của ARN xác định. Quá trình này được gọi là **sự dịch mã**. Trình tự các amino acid trong chuỗi polypeptid xác định cấu trúc ba chiều của nó; cấu trúc ba chiều này quy định chức năng của protein. Thuyết trung tâm khẳng định rằng thông tin chỉ được truyền theo một chiều: từ ADN đến ARN, rồi đến protein. Tuy nhiên, người ta đã tìm thấy trường hợp ngoại lệ của quy luật này ở retrovirus. Ở nhóm các virus này, có enzym phiên mã ngược (reverse transcriptase) có khả năng phiên mã theo chiều từ ARN thành ADN. Hoạt động chức năng của tế bào cũng như của cơ thể sống phụ thuộc vào hoạt động điều hoà của nhiều loại protein. Thông tin di truyền được lưu giữ trong các gen chính là một chương trình hướng dẫn chi tiết về quá trình biểu hiện của các gen, để đảm bảo sự tổng hợp các loại protein cần thiết cho tế bào luôn xảy ra đúng lúc và đúng chỗ.



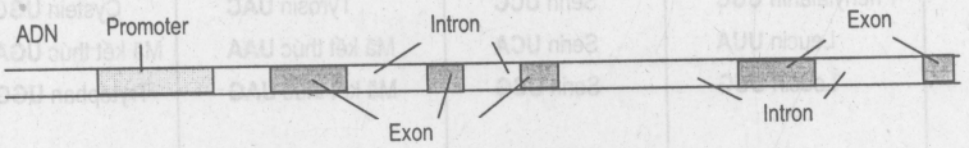
Hình 1.10. **Lý thuyết trung tâm**

Trình tự khởi đầu phiên mã

Sự biểu hiện của gen được điều khiển rất chặt chẽ. Không phải tất cả các gen có trong ADN của tế bào đều được biểu hiện đồng thời. Những gen khác nhau được hoạt hoá biểu hiện vào những thời điểm và ở những tế bào khác nhau. Tất cả các gen được biểu hiện trong một tế bào sẽ xác định đặc tính và chức năng của tế bào đó trong cơ thể. Ví dụ, các gen biểu hiện trong tế bào cơ khác với các gen được biểu hiện trong tế bào máu. Sự biểu hiện của gen được điều khiển bởi một đoạn trình tự ADN đứng trước đoạn trình tự mã hóa được gọi là **trình tự khởi đầu phiên mã** (promoter, còn gọi là trình tự khởi động). Đoạn trình tự khởi động chứa trình tự đặc hiệu được ARN polymerase và các protein đặc biệt gọi là các **yếu tố phiên mã** nhận biết để gắn vào trong quá trình phiên mã của gen. Mức độ biểu hiện của gen trong tế bào được xác định bằng mức độ gắn kết (ái lực) của ARN polymerase và các yếu tố phiên mã với promoter.

Intron và Exon

Một đặc tính rất lạ của các gen ở các sinh vật bậc cao (sinh vật nhân chuẩn) là thông tin di truyền được mã hoá thường bị phân cắt thành nhiều đoạn trình tự ADN cách biệt được gọi là các **exon**. Các exon bị ngăn cách bởi những trình tự không mang thông tin có ích được gọi là các **intron** (*Hình 1.11*). Số lượng các intron trong một gen biến động lớn, từ 0 đến trên 50. Độ dài của các intron và các exon cũng rất biến động, nhưng các intron thường dài hơn và chiếm phần lớn trình tự của gen. Trước khi thông tin trong gen có thể được sử dụng để tổng hợp phân tử protein tương ứng mà nó mã hóa, thì các intron phải bị cắt bỏ khỏi phân tử ARN nhờ quá trình được gọi là **quá trình cắt bỏ** (quá trình hoàn thiện phân tử mARN). Trong quá trình đó, các exon được giữ lại và nối lại với nhau thành một trình tự mã hoá liên tục. Các intron là đặc điểm xuất hiện phổ biến trong hệ gen của các sinh vật bậc cao và thường không thấy có ở vi khuẩn.



Hình 1.11. **Cấu trúc của gen**

Gen giả

Có một số gen giống với các gen khác nhưng trình tự base của chúng có những sai sót làm cho chúng không có khả năng chứa những thông tin sinh học hữu ích. Những gen đó được gọi là những gen giả và là những gen mang những sai sót hoặc đột biến trong trình tự ADN của chúng trong quá trình tiến hoá làm cho thông tin bị lẫn lộn đến mức không còn điều khiển quá trình sinh tổng hợp protein được nữa. Những gen giả đó là những dấu ấn của quá trình tiến hoá. Trải qua tiến hoá, từ những sự biến đổi ban đầu, các base gây mất thông tin được lặp đi lặp lại đến mức, cuối cùng, trình tự base của các gen giả khác hẳn với gen gốc. Ví dụ như các gen globin giả trong các cụm gen globin.

I.C MÃ DI TRUYỀN

Sự biểu hiện

Thông tin di truyền được mã hoá bởi trình tự các nucleotid (các base) và được tổ chức thành các gen. Sự biểu hiện của gen là một thuật ngữ mô tả quá trình tế bào giải mã thông tin di truyền và sử dụng các thông tin đó để tổng hợp nên các phân tử protein cần cho các hoạt động chức năng của tế bào. Trong quá trình biểu hiện, thông tin di truyền được phiên mã từ ADN sang ARN bằng cách tổng hợp một phân tử ARN có trình tự base bổ trợ với trình tự base của mạch ADN khuôn. Đối với mỗi gen, trình tự của ADN tương đồng tuyến tính với trình tự amino acid của chuỗi polypeptid mà nó mã hoá theo cách trình tự base theo chiều 5'→ 3' của mạch ARN xác định trình tự amino acid của chuỗi polypeptid từ đầu amino (đầu N) đến đầu carboxyl (đầu C).

Mã di truyền

Mã di truyền mô tả cách thức mà trình tự các base có thể chuyển đổi thành trình tự các amino acid trong quá trình tổng hợp protein. Trình tự các base của gen được chia thành nhiều đơn vị nhỏ, mỗi đơn vị gồm ba base liên tiếp. Mỗi bộ ba base như vậy được gọi là một **bộ ba** (hay **codon**) và xác định một amino acid tương ứng. Bốn loại base trong phân tử ADN hoặc ARN có thể tổ hợp thành $4^3 = 64$ loại codon, mã hóa cho 20 loại amino acid của protein (*Bảng 1.1*). Vì số các bộ ba lớn,

Bảng 1.1. **Mã di truyền**

Vị trí thứ nhất (đầu 5')	Vị trí thứ hai				Vị trí thứ ba (đầu 3')
	U	C	A	G	
U	Phenylalanin UUU	Serin UCU	Tyrosin UAU	Cystein UGU	U
	Phenylalanin UUC	Serin UCC	Tyrosin UAC	Cystein UGC	C
	Leucin UUA	Serin UCA	Mã kết thúc UAA	Mã kết thúc UGA	A
	Leucin UUG	Serin UCG	Mã kết thúc UAG	Trytophan UGG	G
C	Leucin CUU	Prolin CCU	Histidin CAU	Arginin CGU	U
	Leucin CUC	Prolin CCC	Histidin CAC	Arginin CGC	C
	Leucin CUA	Prolin CCA	Glycin CAA	Arginin CGA	A
	Lecin CUG	Prolin CCG	Glycin CAG	Arginin CGG	G
A	Isoleucin AUU	Threonin ACU	Asparagin AAU	Serin AGU	U
	Isoleucin AUC	Threonin ACC	Asparagin AAC	Serin AGC	C
	Isoleucin AUA	Threonin ACA	Lysin AAA	Arginin AGA	A
	Methionin AUG	Threonin ACG	Lysin AAG	Arginin AGG	G
G	Valin GUU	Alanin GCU	Acid Aspartic GAU	Glycin GGU	U
	Valin GUC	Alanin GCC	Acid Aspartic GAC	Glycin GGC	C
	Valin GUA	Alanin GCA	Acid Glutamic GAA	Glycin GGA	A
	Valin GUG	Alanin GCG	Acid Glutamic GAG	Glycin GGG	G

nên phần lớn các loại amino acid (trừ methionin và tryptophan), đều được mã hoá bởi nhiều hơn một bộ ba. Đặc tính này của mã di truyền được gọi là **tính suy biến** (hoặc **tính dư thừa,** hay **tính thoái hóa** của mã bộ ba). Các bộ ba cùng quy định một amino acid được gọi là các **bộ ba đồng nghĩa** và chúng gần giống nhau. Ví dụ ACU, ACC, ACA và ACG cùng xác định amino acid threonin. Sự khác biệt giữa các base đồng nghĩa thường chỉ ở vị trí base thứ ba, vị trí này được gọi là **vị trí linh hoạt**. Tính suy biến của mã di truyền giảm thiểu được tác động của đột biến. Nghĩa là sự thay đổi trình tự base trên gen đôi khi không làm thay đổi amino acid trên phân tử protein và có thể tránh được những hậu quả có hại đối với chức năng của protein. Trong số 64 bộ ba, chỉ có 61 bộ ba mã hoá cho các amino acid. Ba bộ ba còn lại là UAG, UGA và UAA không mã hóa cho amino acid nào mà là tín hiệu kết thúc việc tổng hợp protein (dịch mã) và được gọi là các **bộ ba kết thúc** hay **stop codon**. Bộ ba quy định methionin là AUG cũng đồng thời là tín hiệu bắt đầu quá trình dịch mã và được gọi là **bộ ba khởi đầu**. Như vậy, tất cả các chuỗi polypeptid đều bắt đầu bằng methionin, cho dù, đôi khi amino acid này sau đó bị loại bỏ trong quá trình hoàn thiện protein.

Khung đọc

Ngoài việc quy định điểm bắt đầu quá trình tổng hợp protein, bộ ba mã khởi đầu (AUG) còn xác định khung đọc của trình tự ARN. Có thể có ba bộ ba cho bất kỳ một trình tự base nào, phụ thuộc vào base nào được chọn làm base bắt đầu của codon. Thực tế trong quá trình tổng hợp protein, thường chỉ có một khung đọc được sử dụng. Còn hai khung đọc kia thường chứa một số bộ ba kết thúc ngăn cản chúng được sử dụng để tổng hợp trực tiếp nên phân tử protein (*Hình 1.12*).

Khung đọc 1. 5' -	AUG	ACU	AAG	AGA	UCC	GG - 3'
	Met	Thr	Lys	Arg	Ser	
Khung đọc 2. 5' -	A UGA	CUA	AGA	GAU	CCG	G - 3'
	Stop	Leu	Arg	Asp	Pro	
Khung đọc 3. 5' -	AU GAC	UAA	GAG	AUC	CGG	- 3'
	Asp	Stop	Glu	Ile	Arg	

Hình 1.12. **Mỗi trình tự ADN đều có thể đọc theo ba khung đọc khác nhau, phụ thuộc vào base nào được chọn làm base khởi đầu**

Đoạn trình tự nằm giữa một bộ ba khởi đầu và một bộ ba kết thúc tương ứng cùng khung đọc được gọi là **khung đọc mở** (ORF = *o*pen *r*eading *f*rame). Đặc điểm này được dùng để xác định các trình tự ADN mã hoá protein trong các dự án giải mã hệ gen.

Tính vạn năng của mã di truyền

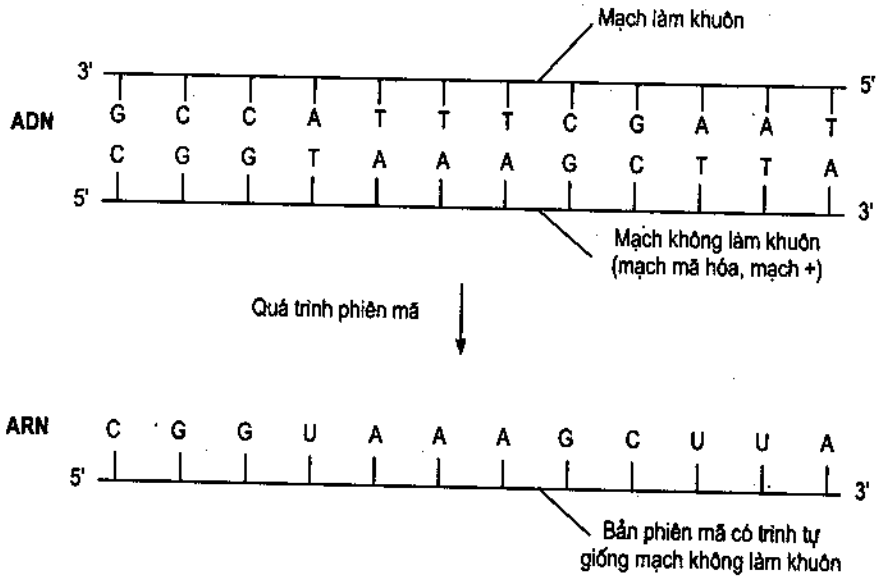
Ban đầu, người ta tin rằng mã di truyền là vạn năng. Nghĩa là ở mọi sinh vật, các codon giống nhau đều quy định những amino acid như nhau. Tuy vậy, thực tế cho thấy có một số trường hợp ngoại lệ hiếm gặp. Ví dụ, ở hệ gen ty thể có sự khác biệt về bộ ba khởi đầu và bộ ba kết thúc. Cụ thể, AUG bình thường là bộ ba kết thúc, thì ở ty thể nó lại mã hoá cho tryptophan; AGA và AGG bình thường quy định arginin, ở ty thể lại có vai trò là các bộ ba kết thúc; AUA bình thường mã hóa cho isoleucin thì ở ty thể lại xác định methionin. Người ta cho rằng những thay đổi này có thể tồn tại được là nhờ ty thể là một hệ thống kín. Ngoài hệ gen ty thể, một số trường hợp ngoại lệ khác cũng được tìm thấy ở một số sinh vật đơn bào. Ví dụ ở một số động vật nguyên sinh, các bộ ba UAA và UAG bình thường là các bộ ba kết thúc thì lại mã hoá cho acid glutamic.

I.D PHIÊN MÃ

Phiên mã

Trong quá trình này, một bản sao trình tự của gen được tạo ra và đây là giai đoạn đầu của quá trình biểu hiện của gen. ARN được tạo ra nhờ enzym ARN polymerase trên mạch khuôn ADN. Một mạch của ADN là mạch khuôn còn mạch kia là mạch không làm khuôn. ARN chính là bản sao của mạch không làm khuôn vì nó được tổng hợp từ mạch làm khuôn theo nguyên tắc bổ sung (*Hình 1.13*). Ta thường lấy trình tự của mạch không làm khuôn làm trình tự của gen. Người ta còn dùng các tên gọi khác để chỉ mạch không làm khuôn, chẳng hạn như mạch có nghĩa (+) hay mạch mang mã. Phân tử ARN được tổng hợp gọi là bản phiên mã.

Trong quá trình phiên mã, ARN được tổng hợp nên bằng cách polyme hoá các ribonucleotid triphosphat (ATP, UTP, GTP, CTP). Đầu 3'-OH của ribonucleotid này tương tác với đầu 5'-phosphat của ribonucleotid kia để tạo ra một liên kết phosphodieste. Trật tự các ribonucleotid thêm vào phân tử ARN đang tổng hợp do trình tự base của ADN khuôn quy định. Nucleotid mới được thêm vào đầu 3' tự do của chuỗi. Bản phiên mã được tổng hợp theo chiều 5' → 3' lấy sợi ADN có chiều 3' → 5' làm khuôn.



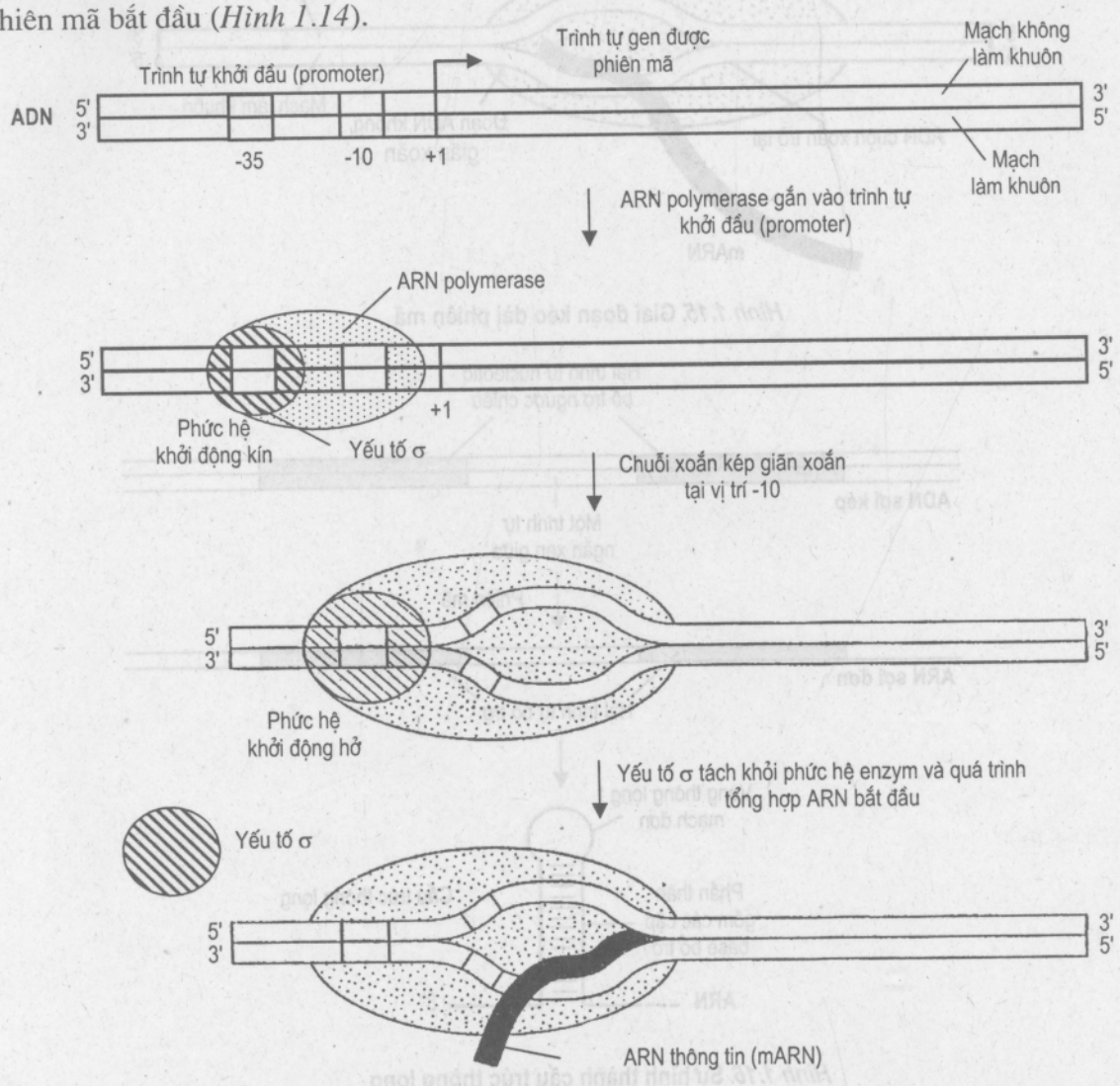
Hình 1.13. **Quá trình phiên mã mạch ADN làm khuôn tạo ra bản phiên mã ARN có trình tự giống với mạch ADN không làm khuôn**

Phiên mã ở sinh vật nhân sơ

Ở sinh vật nhân sơ, quá trình phiên mã diễn ra theo ba pha là khởi đầu, kéo dài và kết thúc phiên mã. ARN được tổng hợp nhờ enzym ARN polymerase có chứa nhiều tiểu đơn vị là các chuỗi polypeptid. Ở *E.coli*, ARN polymerase có 5 tiểu đơn vị là hai chuỗi α, một chuỗi β, một chuỗi β' và tiểu đơn vị σ ($\alpha_2\beta'\beta\sigma$). Dạng này được gọi là holoenzym. Tiểu đơn vị σ có thể tách khỏi các tiểu đơn vị khác để lại phần được gọi là enzym lõi. Hai dạng này của ARN polymerase có những chức năng khác nhau trong quá trình phiên mã.

Sự khởi đầu phiên mã

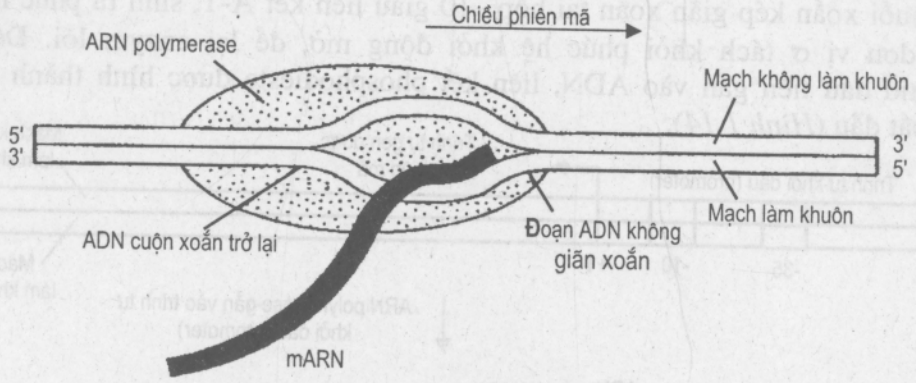
Quá trình phiên mã không thể bắt đầu một cách ngẫu nhiên mà phải tại điểm bắt đầu của gen. Tín hiệu khởi đầu phiên mã nằm ở đoạn trình tự khởi động (promoter, còn gọi là *trình tự khởi đầu phiên mã*); đoạn này nằm ở phía trước đoạn trình tự được phiên mã của gen. Đoạn khởi động có chứa trình tự ADN đặc hiệu hoạt động như những điểm gắn của ARN polymerase. Ở *E. coli*, có hai trình tự được ARN polymerase nhận biết được gọi là đoạn trình tự –10 và đoạn trình tự –35. Các trình tự này có thể khác nhau ở những đoạn khởi động của các gen khác nhau nhưng tất cả đều tuân theo một khuôn mẫu chung được gọi là trình tự liên ứng. Tiểu đơn vị σ của ARN polymerase chịu trách nhiệm nhận biết và gắn vào đoạn khởi động, có lẽ tại trình tự –35. Nếu không có tiểu đơn vị σ, ARN polymease vẫn gắn được vào ADN nhưng một cách ngẫu nhiên. Khi enzym gắn được vào đoạn khởi động, thì đầu tiên nó hình thành một phức hệ khởi động kín, trong đó đoạn ADN khởi động vẫn ở dạng xoắn kép. Enzym ARN polymerase phủ một vùng khoảng 60 cặp base kể cả hộp –10 và –35. Để quá trình phiên mã bắt đầu, chuỗi xoắn kép giãn xoắn tại hộp –10 giàu liên kết A-T, sinh ra phức hệ khởi động mở. Tiểu đơn vị σ tách khỏi phức hệ khởi động mở, để lại enzym lõi. Đồng thời hai ribonucleotid đầu tiên gắn vào ADN, liên kết phosphodieste được hình thành và quá trình phiên mã bắt đầu (*Hình 1.14*).



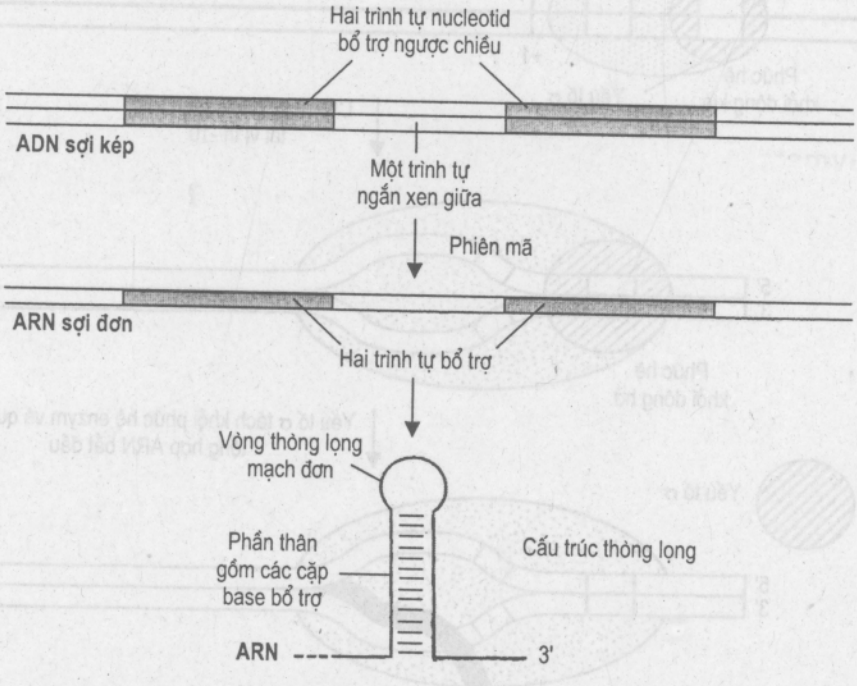
Hình 1.14. **Sự khởi đầu phiên mã ở sinh vật nhân sơ**

Kéo dài phiên mã

ARN dịch chuyển dọc theo phân tử ADN giãn xoắn. Enzym này gắn thêm các ribonucleotid theo trật tự do các base trên mạch ADN khuôn quy định vào đầu 3' của phân tử ARN đang dài ra. Ở hầu hết các trường hợp, có một **đoạn trình tự dẫn đầu**, có độ dài dao động, được phiên mã trước đoạn trình tự mã hoá của gen. Tương tự, ở cuối đoạn trình tự mã hoá còn có một đoạn trình tự không mang mã được gọi là đoạn trình tự theo sau cũng được phiên mã trước khi kết thúc phiên mã. Trong quá trình phiên mã, ở bất kỳ thời điểm nào cũng chỉ có một phần nhỏ chuỗi xoắn kép giãn xoắn. Vùng giãn xoắn này chứa sợi ARN đang được tổng hợp ghép đôi base với sợi ADN khuôn trong phạm vi 12 - 17 base. Đoạn giãn xoắn phải nhỏ vì sự giãn xoắn ở một vùng làm căng xoắn ở vùng liền kề và điều đó gây căng cứng phân tử ADN. Để khắc phục điều này, phân tử ARN vừa được tổng hợp vừa được giải phóng ra để chuỗi xoắn kép tái tạo và cuộn xoắn trở lại (*Hình 1.15*).



Hình 1.15. **Giai đoạn kéo dài phiên mã**



Hình 1.16. **Sự hình thành cấu trúc thòng lọng**

Kết thúc phiên mã

Phiên mã kết thúc tại những vị trí đặc thù nằm sau trình tự mã hoá. Ở *E. coli*, phiên mã kết thúc tại đoạn trình tự được gọi là đoạn đối xứng. Đoạn này đối xứng qua phần giữa, nghĩa là nửa đầu của đoạn này có trình tự base bổ trợ chính xác với nửa sau. Trên phân tử ARN, đặc điểm này cho phép nửa đầu của đoạn trình tự này bắt cặp base với nửa sau, tạo ra cấu trúc được gọi là cấu trúc thòng lọng (*Hình 1.16*).

Cấu trúc thòng lọng là dấu hiệu kết thúc phiên mã. Trong một số trường hợp, trên ADN, sau đoạn trình tự thòng lọng là khoảng 5 - 10 base A, tạo liên kết poly(A-U) yếu với phân tử ARN mới được tổng hợp. Người ta cho rằng, ARN polymerase dừng lại ngay sau thòng lọng và các cặp base poly(A-U) liên kết yếu đứt ra, làm cho bản phiên mã tách ra khỏi mạch khuôn. Trong những trường hợp khác, không có đoạn base poly(A) mà thay vào đó có một cơ chế khác: đó là protein Rho (ρ) gắn vào trình tự kết thúc phiên mã, phá vỡ liên kết giữa các base của mạch khuôn với bản phiên mã khi enzym polymerase dừng lại sau thòng lọng. Quá trình kết thúc phiên mã còn bao gồm việc giải phóng bản phiên mã và enzym lõi để enzym này có thể gắn với một tiểu đơn vị σ mới và tham gia một hành trình phiên mã khác.

Phiên mã ở sinh vật nhân chuẩn

Tương tự như ở sinh vật nhân sơ, song giai đoạn khởi đầu phiên mã phức tạp hơn; giai đoạn kết thúc phiên mã không hình thành cấu trúc thòng lọng và sự phiên mã do ba enzym tiến hành (ARN polymerase I, II và III), mỗi enzym phiên mã một bộ các gen khác nhau.

ARN polymerase I

Enzym này phiên mã các gen mã hoá cho ba trong số bốn ARN ribosom (18S, 28S và 5,8S). Đoạn khởi động mà ARN polymerase I nhận biết được có hai trình tự thành phần quan trọng: thành phần lõi phủ lên điểm bắt đầu phiên mã và đoạn trình tự kiểm soát ngược dòng nằm cách điểm bắt đầu phiên mã khoảng 100 cặp base. Tín hiệu kết thúc gồm một trình tự liên ứng 18 cặp base nằm sau điểm kết thúc của gen khoảng 600 cặp base.

ARN polymerase II

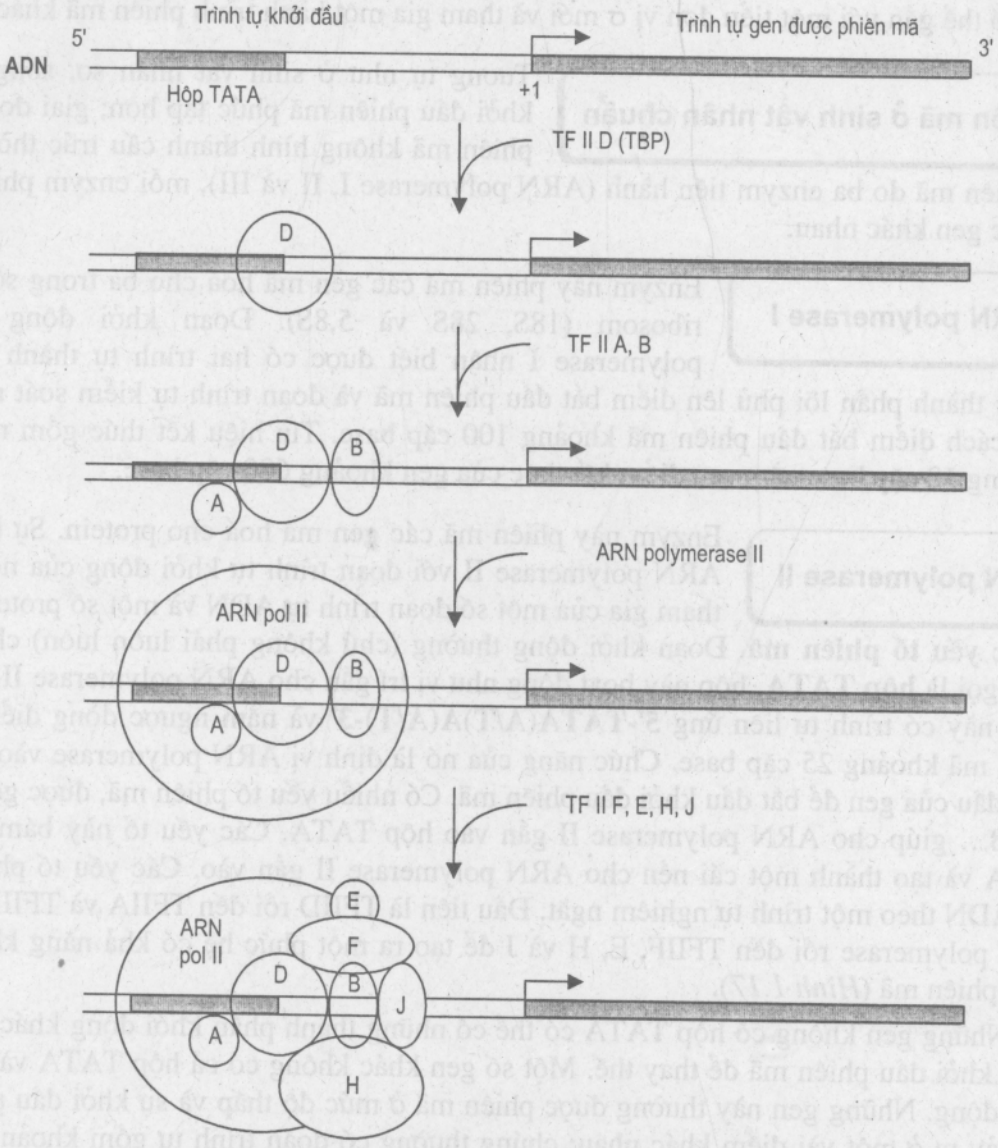
Enzym này phiên mã các gen mã hoá cho protein. Sự liên kết của ARN polymerase II với đoạn trình tự khởi động của nó cần có sự tham gia của một số đoạn trình tự ADN và một số protein được gọi là các **yếu tố phiên mã**. Đoạn khởi động thường (chứ không phải luôn luôn) chứa trình tự được gọi là **hộp TATA**, hộp này hoạt động như vị trí gắn cho ARN polymerase II. Đoạn khởi động này có trình tự liên ứng **5'-TATA(A/T)A(A/T)-3'** và nằm ngược dòng điểm khởi đầu phiên mã khoảng 25 cặp base. Chức năng của nó là định vị ARN polymerase vào đúng điểm khởi đầu của gen để bắt đầu khởi đầu phiên mã. Có nhiều yếu tố phiên mã, được gọi là TFIIA, TFIIB... giúp cho ARN polymerase II gắn vào hộp TATA. Các yếu tố này bám quanh hộp TATA và tạo thành một cái nền cho ARN polymerase II gắn vào. Các yếu tố phiên mã gắn vào ADN theo một trình tự nghiêm ngặt. Đầu tiên là TFIID rồi đến TFIIA và TFIIB. Sau đó là ARN polymerase rồi đến TFIIF, E, H và J để tạo ra một phức hệ có khả năng khởi đầu quá trình phiên mã (*Hình 1.17*).

Những gen không có hộp TATA có thể có những thành phần khởi động khác nằm quanh điểm khởi đầu phiên mã để thay thế. Một số gen khác không có cả hộp TATA và thành phần khởi động. Những gen này thường được phiên mã ở mức độ thấp và sự khởi đầu phiên mã có thể xảy ra ở một vài điểm khác nhau; chúng thường có đoạn trình tự gồm khoảng 100 - 200 cặp base giàu GC nằm ngược dòng đoạn khởi đầu phiên mã.

Một số thành phần khởi động khác có những trình tự liên ứng đặc trưng ảnh hưởng đến sự phiên mã. Các trình tự này hoạt động như những vị trí gắn kết cho các yếu tố phiên mã khác làm nhiệm vụ điều hoà phiên mã bằng cách kích thích hoặc ức chế sự khởi đầu phiên mã. Ví dụ như hộp CCAAAT (còn gọi là hộp CAT) nằm ngược dòng với hộp TATA ở một số gen. Nhiều gen còn chứa các trình tự được gọi là các **gen tăng cường** có tác dụng kích thích mạnh quá trình phiên mã. Các gen này thường nằm bên ngoài đoạn trình tự khởi động, xa gen mà nó ảnh hưởng tới và hoạt động độc lập. Ngoài ra, còn có các **gen bất hoạt** ức chế sự phiên mã.

Cơ chế để ARN polymerase II kết thúc phiên mã còn chưa rõ lắm. Rất khó xác định tín hiệu kết thúc vì đầu 3' thường bị cắt ngắn đi sau khi nó được phiên mã. Có thể là không tồn tại tín hiệu kết thúc phiên mã duy nhất nào, cũng có thể sự tách ra của một yếu tố phiên mã ở một vị trí nào đó phá vỡ sự bền vững của phức hệ phiên mã làm cho ARN polymerase, sau đó, rời khỏi sợi khuôn.

Sự khởi đầu phiên mã của ARN polymerase I và III cũng giống như II, cũng bao gồm những đoạn trình tự khởi động làm vị trí gắn cho ARN polymerase và các yếu tố phiên mã (TF-I cho ARN polymerase I và TFIII A-C cho ARN polymerase III).



Hình 1.17. **Quá trình gắn của enzym ARN polymerase II và các yếu tố phiên mã vào trình tự khởi đầu phiên mã (promoter) ở sinh vật nhân chuẩn**

ARN polymerase III

Enzym này phiên mã một nhóm những gen ngắn mã hoá cho các ARN vận chuyển và ARN ribosom 5S. Trình tự khởi đầu mà ARN polymease III nhận biết có một đặc điểm là nằm sau điểm khởi đầu phiên mã và trong đoạn trình tự phiên mã của gen. Những đoạn trình tự ấy được gọi là vùng kiểm soát trong ICR (*i*nternal *c*ontrol *r*egion) và nằm cách điểm khởi đầu phiên mã khoảng 100 cặp base. Gen rARN 5S có đoạn trình tự được gọi là hộp C là vị trí gắn kết của các yếu tố phiên mã và enzym ARN polymerase III. Một đoạn trình tự thứ hai nằm ngược dòng được gọi là hộp A cũng có vai trò quan trọng. Đối với các gen tARN, các ICR chứa hai đoạn trình tự có tính bảo thủ cao gọi là hộp A và hộp B, là những vị trí gắn kết của các yếu tố phiên mã và enzym ARN polymerase III. Sự kết thúc phiên mã của enzym ARN polymerase III xảy ra tại đoạn trình tự gồm một dãy base adenin (poly(A)) xuất hiện sau điểm kết thúc của gen.

I.E ARN VẬN CHUYỂN (tARN)

Vai trò của tARN trong quá trình dịch mã

Mọi tế bào đều có ba loại ARN, đó là ARN vận chuyển (tARN), ARN ribosom (rARN) và ARN thông tin (mARN). Tất cả các phân tử ARN đều được phiên mã từ ADN. ARN vận chuyển và ARN ribosom là một bộ phận của bộ máy tổng hợp protein, còn ARN thông tin là khuôn để tổng hợp protein trong quá trình dịch mã.

Các phân tử tARN có kích thước nhỏ làm nhiệm vụ "phiên dịch" trình tự nucleotid của mARN thành trình tự amino acid của chuỗi polypeptid tương ứng. Trong tế bào có nhiều loại phân tử tARN. Mỗi loại có thể gắn với một loại amino acid nhất định. Đồng thời, mỗi loại tARN có thể nhận biết một bộ ba nucleotid đặc trưng trên phân tử mARN. Bằng cách như vậy, phân tử tARN có thể đặt aminoacid vào đúng vị trí trên chuỗi polypeptid tương ứng với trình tự nucleotid của phân tử mARN.

Cấu trúc tARN

Một phân tử tARN thường chứa khoảng 74 - 95 nucleotid. Trong phân tử, có những đoạn trình tự nucleotid bổ trợ có khả năng kết cặp tạo cho phân tử ARN có hình chiếc lá với nhiều thuỳ. Đây là cấu trúc đặc trưng cho tARN (*Hình 1.18a*). Các thùy bao gồm:

– Thùy nhận amino acid được hình thành do sự kết cặp của các base ở hai đầu 5' và 3' của phân tử tARN. Trình tự CCA ở tận cùng đầu 3' không kết cặp và là vị trí gắn của amino acid.

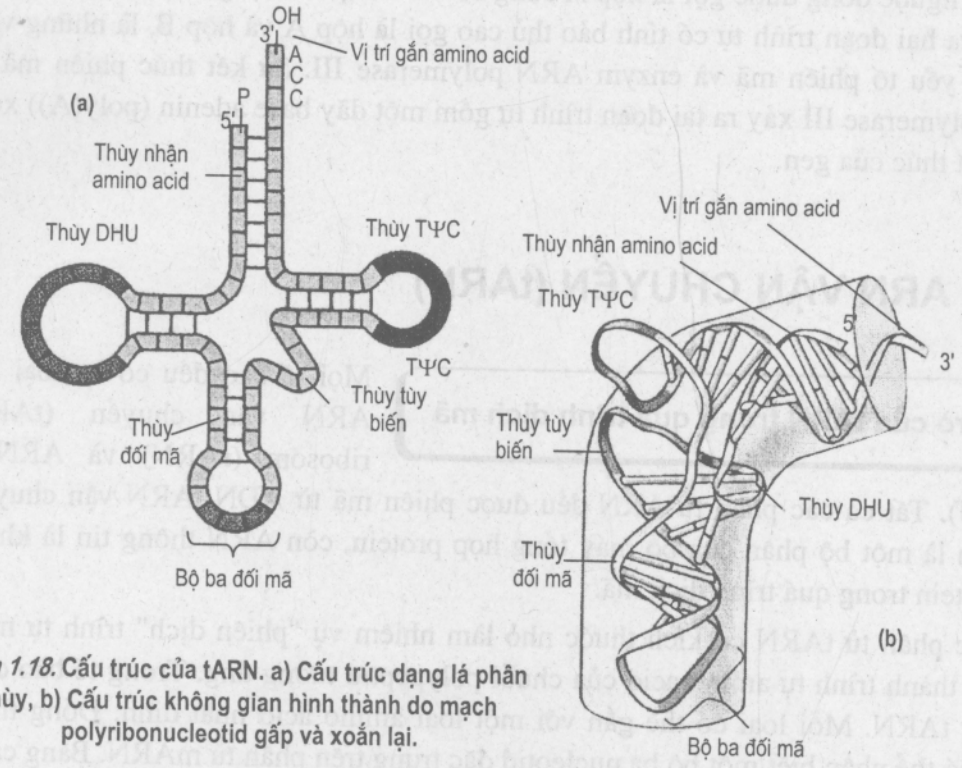
– Thùy D hay còn gọi là thuỳ DHU là thùy có cấu trúc thòng lọng chứa dihydrouracil (DHU), một loại nucleotid pyrimidin khác thường.

– Thùy đối mã chịu trách nhiệm nhận biết bộ ba mã hoá trên mARN.

– Thùy tùy biến hay còn gọi là thùy biến động chỉ có ở tARN. Thùy này có thể có kích thước nhỏ (chỉ 2-3 nucleotid, như ở nhóm tARN I) hoặc lớn (tới 13 – 21 nucleotid với 5 cặp base kết cặp bổ trợ như ở nhóm tARN II).

– Thùy TφC chứa trình tự TφC, trong đó φ là nucleotid bị cải biến được gọi là uracil giả (pseudouracil).

Khi so sánh các loại phân tử tARN khác nhau, người ta thấy sự xuất hiện của nhiều đoạn trình tự có tính bảo thủ cao; trong những đoạn trình tự này, các nucleotid là giống hệt nhau ở tất cả mọi loại tARN. Ở một số vị trí khác luôn là purin hoặc pyrimidin, được gọi là những vị trí bán cố định. Trên những phần còn lại là các trình tự nucleotid có mức độ biến động cao giữa các loại phân tử tARN khác nhau.



Hình 1.18. **Cấu trúc của tARN. a) Cấu trúc dạng lá phân thùy, b) Cấu trúc không gian hình thành do mạch polyribonucleotid gấp và xoắn lại.**

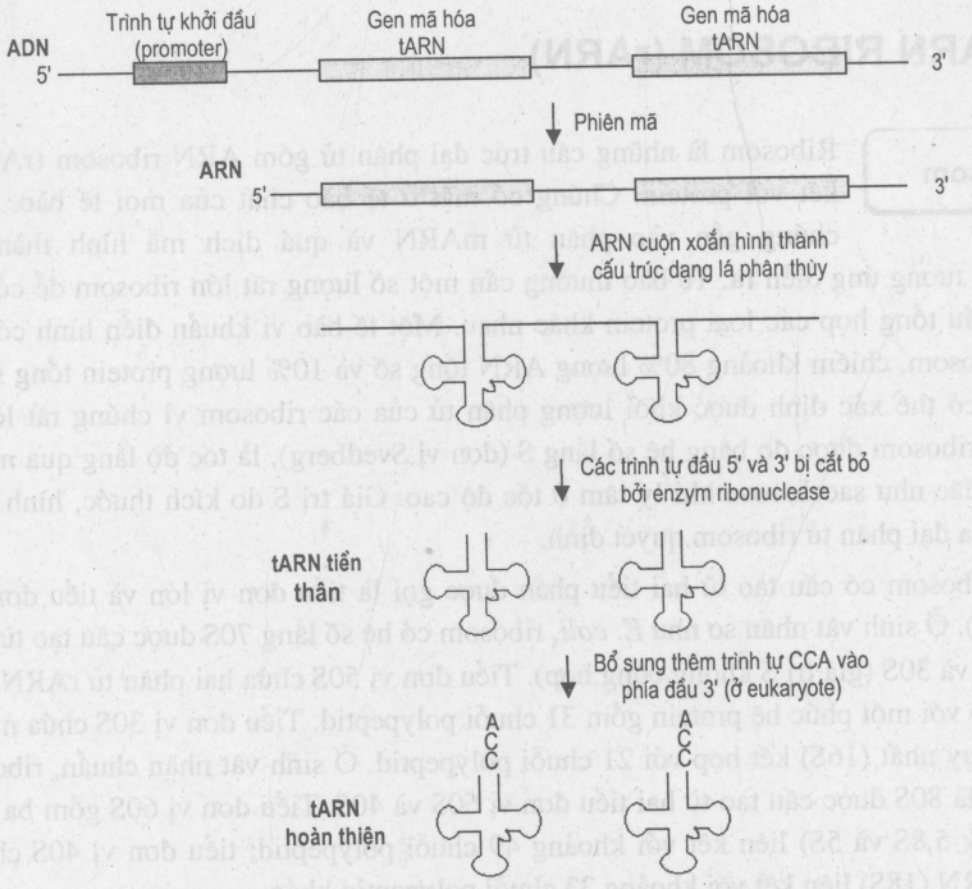
Cấu trúc lá có thùy như nêu trên là hình ảnh hai chiều của tARN. Hình ảnh ba chiều (3D) của các phân tử tARN có thể được xác định nhờ nhiễu xạ tia X (*Hình 1.18b*). Ngoài những trình tự base kết cặp thấy được trên cấu trúc hai chiều, hình ảnh ba chiều còn cho thấy một số base kết cặp khác nữa. Rất nhiều base kết cặp ở dạng cố định hoặc bán cố định. Trên hình ảnh 3D, thùy nhận amino acid vẫn nằm đối diện với thùy đối mã, đảm bảo chức năng của nó trong quá trình dịch mã.

Tổng hợp và hoàn thiện

Các phân tử tARN được tổng hợp bằng cách phiên mã các gen mã hóa cho các loại phân tử tARN nhờ enzym ARN polymerase III. Các gen này có nhiều bản sao trong tế bào, đặc biệt đúng ở các tế bào nhân chuẩn. Điều này phản ánh nhu cầu của tế bào cần lượng lớn các phân tử tARN.

Các phân tử tARN ban đầu được hình thành dưới dạng các phân tử tARN tiền thân (gọi tắt là tiền-tARN) rồi được cải biến thành các phân tử tARN hoàn thiện (*Hình 1.19*). Một số

gen mã hóa tARN khác nhau có thể được phiên mã cùng lúc để hình thành một phân tử tARN tiền thân chung, rồi sau đó được hoàn chỉnh nhờ enzym ribonuclease cắt tại vị trí các đầu 5' và 3' của mỗi loại tARN riêng biệt. Ở sinh vật nhân sơ, quá trình hoàn thiện tARN được thực hiện qua nhiều bước với sự tham gia của các enzym ribonuclease D và P (viết tắt là ARNaseD và ARNaseP). ARNaseP có mặt ở cả sinh vật nhân sơ và nhân chuẩn. Một đặc điểm đặc biệt là enzym này có thành phần ARN có hoạt tính xúc tác được gọi là ribozym. Các phân tử tARN của sinh vật nhân chuẩn khác với các tARN của sinh vật nhân sơ ở chỗ: nhiều phân tử tARN được phiên mã mang một trình tự intron ngắn. Trình tự này chỉ bị loại đi trong quá trình hoàn chỉnh tARN. Trình tự CCA ở tận cùng đầu 3' của tất cả các phân tử tARN là điểm gắn của amino acid. Ở sinh vật nhân chuẩn, CCA không có trên trình tự ADN của gen mã hóa cho tARN, nhưng trong quá trình hoàn thiện tARN sau đó, nó được bổ sung thêm vào nhờ enzym tARN nucleotidyl transferase. Ở sinh vật nhân sơ, CCA có sẵn trên trình tự gen mã hoá, nhưng đôi khi bị ARNaseD cắt bỏ rồi được khôi phục lại nhờ hoạt động của enzym nucleotidyl transferase cũng có ở sinh vật nhân sơ.



Hình 1.19. **Quá trình phiên mã và hoàn thiện các phân tử tARN**

Sửa đổi các nucleotid

Các phân tử tARN có chứa những nucleotid khác thường được hình thành sau quá trình phiên mã nhờ quá trình cải biến (sửa đổi) về mặt hoá học. Những dạng cải biến phổ biến nhất gồm có:

– Methyl hoá đường ribose. Ví dụ guanosin bị methyl hoá thành 7-methylguanosin (7-mG).

– Tái sắp xếp lại các base bao gồm trao đổi vị trí các nguyên tử trong vòng purin và pyrimidin. Ví dụ: uridin chuyển thành uridin giả.

– Biến tính liên kết đôi. Ví dụ, uridin thành dihydrouridin (DHU).

– Loại (khử) amin. Quá trình này bao gồm sự loại bỏ nhóm amin khỏi các base. Ví dụ guanosin bị khử amin thành inosin.

– Thay thế bằng nguyên tử lưu huỳnh (S). Ví dụ, nguyên tử oxy của uridin bị thay bằng nguyên tử lưu huỳnh tạo thành 4-thiouridin.

– Bổ sung các nhóm lớn. Ví dụ, guanosin chuyển thành queosine.

Người ta đã xác định được trên 50 kiểu cải biến khác nhau. Mỗi kiểu do một enzym sửa đổi tARN thực hiện. Nguyên nhân của những cải biến này còn chưa được biết đầy đủ. Tuy vậy, trong một số trường hợp, người ta thấy những cải biến này có vai trò đối với những nucleotid nằm bên trong thùy đối mã.

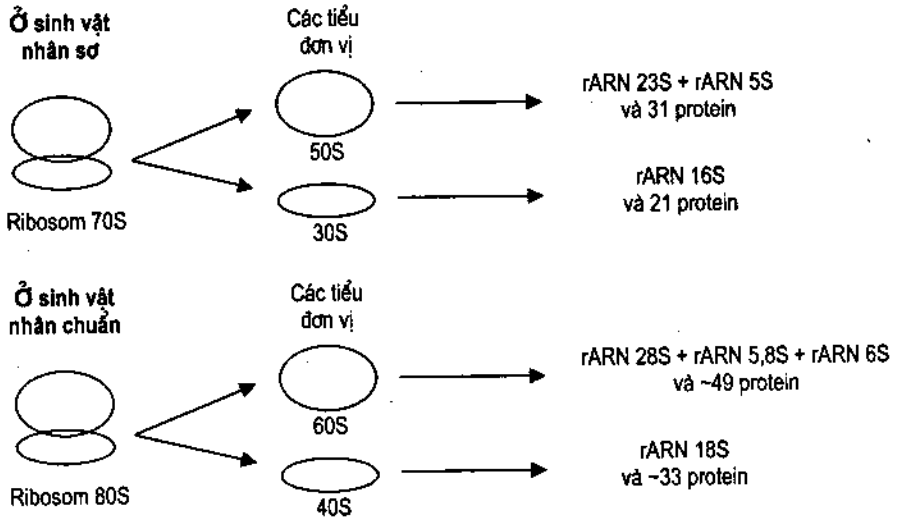
I.F ARN RIBOSOM (rARN)

Ribosom

Ribosom là những cấu trúc đại phân tử gồm ARN ribosom (rARN) gắn kết với protein. Chúng có mặt ở tế bào chất của mọi tế bào. Tại đây, chúng gắn vào phân tử mARN và quá dịch mã hình thành chuỗi polypeptid tương ứng diễn ra. Tế bào thường cần một số lượng rất lớn ribosom để có thể đáp ứng nhu cầu tổng hợp các loại protein khác nhau. Một tế bào vi khuẩn điển hình có khoảng 20.000 ribosom, chiếm khoảng 80% lượng ARN tổng số và 10% lượng protein tổng số của tế bào. Khó có thể xác định được khối lượng phân tử của các ribosom vì chúng rất lớn. Kích thước của ribosom được đo bằng hệ số lắng S (đơn vị Svedberg), là tốc độ lắng qua một dung dịch đậm đặc như saccharose khi ly tâm ở tốc độ cao. Giá trị S do kích thước, hình dạng và cấu trúc của đại phân tử ribosom quyết định.

Mỗi ribosom có cấu tạo từ hai tiểu phần được gọi là tiểu đơn vị lớn và tiểu đơn vị nhỏ (*Hình 1.20*). Ở sinh vật nhân sơ như *E. coli*, ribosom có hệ số lắng 70S được cấu tạo từ hai tiểu đơn vị 50S và 30S (giá trị S không cộng hợp). Tiểu đơn vị 50S chứa hai phân tử rARN (23S và 5S) kết hợp với một phức hệ protein gồm 31 chuỗi polypeptid. Tiểu đơn vị 30S chứa một phân tử rARN duy nhất (16S) kết hợp với 21 chuỗi polypeptid. Ở sinh vật nhân chuẩn, ribosom có hệ số lắng là 80S được cấu tạo từ hai tiểu đơn vị 60S và 40S. Tiểu đơn vị 60S gồm ba phân tử rARN (28S, 5,8S và 5S) liên kết với khoảng 49 chuỗi polypeptid; tiểu đơn vị 40S chứa một phân tử rARN (18S) liên kết với khoảng 33 chuỗi polypeptid khác.

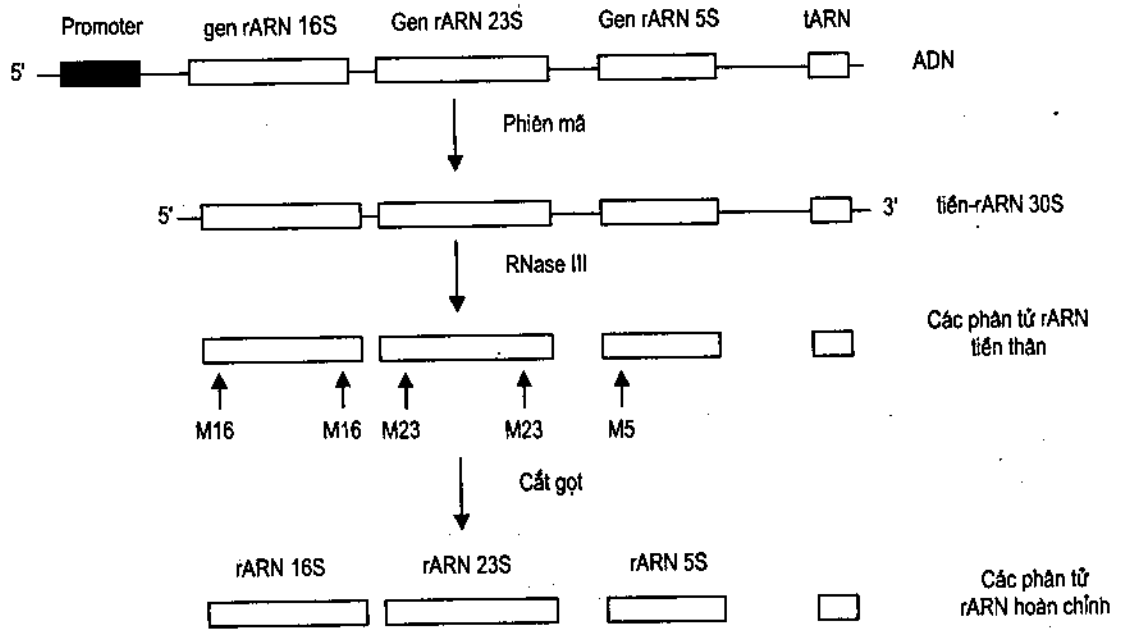
Trong ribosom, các phân tử rARN có cấu trúc ba chiều đặc trưng được xác định nhờ các liên kết giữa các base nitơ bổ trợ bên trong và giữa các phân tử rARN với nhau. Người ta cho rằng các phân tử rARN tạo thành bộ khung, trên đó các phân tử protein có vai trò hoạt động chức năng gắn vào. Một số phân tử ARN có hoạt tính enzym và được gọi là các ribozym và có thể chúng cũng tham gia vào hoạt động chức năng của ribosom.



Hình 1.20. **Thành phần cấu tạo điển hình của ribosom ở sinh vật nhân sơ (prokaryote) và sinh vật nhân chuẩn (eukaryote)**

Phiên mã và hoàn thiện các rARN ở sinh vật nhân sơ

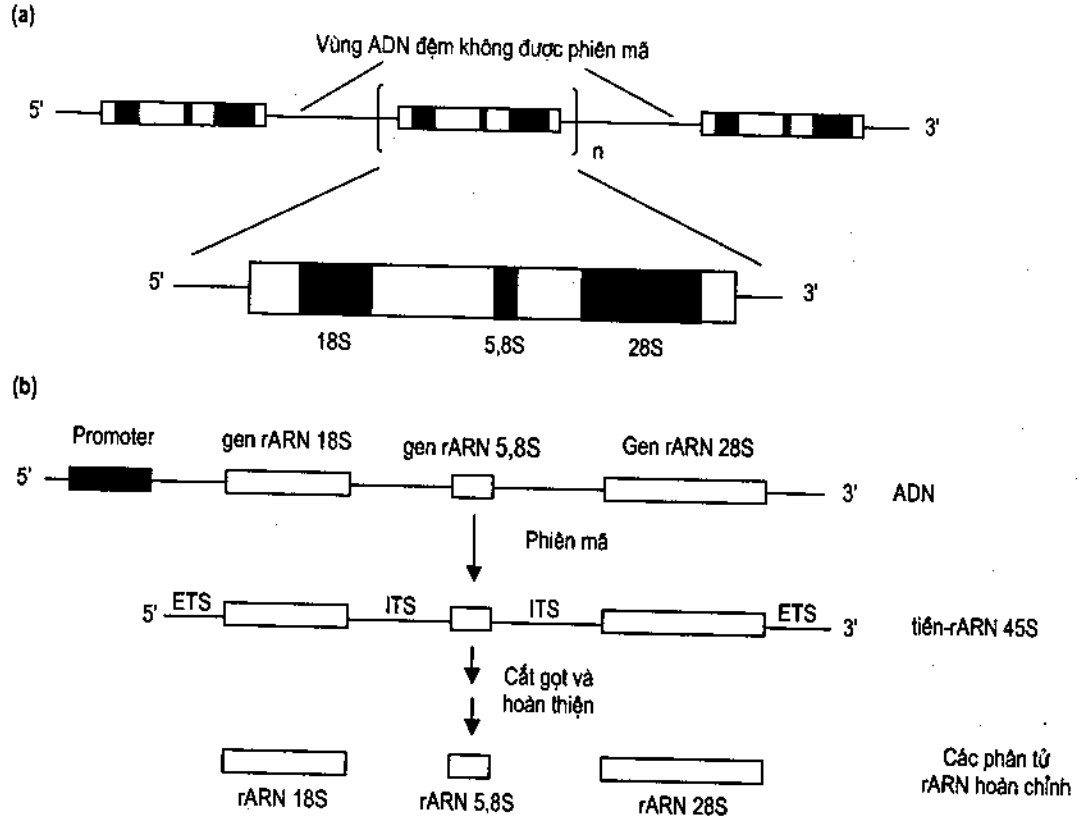
Tế bào thường có một số lượng lớn ribosom. Chúng cũng phải nhân lên cùng với quá trình phân chia tế bào. Do vậy, tế bào cần một số lượng lớn các loại rARN. Các loại rARN này được phiên mã từ các gen mã hóa rARN nhờ enzym ARN polymerase. Để đảm bảo tế bào có đúng số lượng mỗi loại rARN cần thiết, các ribosom được phiên mã cùng lúc từ một gen duy nhất.



Hình 1.21. **Quá trình phiên mã và hoàn thiện các ARN ribosom (rARN) ở *E. coli***

Đáng chú ý là gen này tồn tại trong hệ gen với nhiều bản sao. Ở sinh vật nhân sơ, như *E. coli*, người ta đã xác định có 7 gen rARN phân tán suốt hệ gen. Mỗi gen chỉ có một

bản sao cho các loại rARN 16S, 23S và 5S nằm kế tiếp nhau trong gen. Thêm vào đó, trên mỗi gen có từ 1 đến 4 trình tự mã hóa tARN. Khi gen đó được phiên mã, nó hình thành một phân tử ARN duy nhất được gọi là **tiền-ARN** (30S). Phân tử này sau đó được hoàn thiện để tạo ra các loại rARN và các tARN tương ứng (*Hình 1.21*). Quá trình hoàn thiện gồm nhiều bước. Trước tiên, sau khi phiên mã, phân tử ARN cuộn lại và các trình tự base bổ trợ ghép đôi với nhau tạo ra nhiều cấu trúc thòng lọng. Các protein ribosom gắn vào phân tử ARN đã cuộn xoắn. Tại giai đoạn này, một số base trên phân tử ARN được gắn thêm nhóm methyl (gọi là sự methyl hóa). Cuối cùng, phân tử ARN bị cắt ở những vị trí đặc hiệu bởi enzym ARNase III để giải phóng các phân tử rARN 5S, 23S và 16S. Bước hoàn thiện tiếp theo là cắt bỏ các trình tự ở đầu 3' và 5' nhờ các enzym ribonuclease khác có tên gọi là M5, M16 và M23 để thành các phân tử rARN hoàn chỉnh.



Hình 1.22. **(a) Tổ chức gen rARN ở eukaryote. (b) Phiên mã và hoàn thiện rARN ở eukaryote**

Phiên mã và hoàn thiện các rARN ở sinh vật nhân chuẩn

Ở sinh vật nhân chuẩn, các đoạn trình tự mã hóa cho các phân tử rARN 28S, 18S và 5,8S cũng cùng nằm trên một gen. Gen này tồn tại thành nhiều bản sao phân cách nhau bằng những vùng không mã hóa (*Hình 1.22a*). Ở người có khoảng 200 gen phân bố thành 5 cụm gen, mỗi cụm gồm khoảng 40 gen được phân bố trên các nhiễm sắc thể khác nhau. Các gen này được phiên mã nhờ enzym **ARN polymerase I** trong **nhân con** của tế bào thành phân tử tiền-rARN (45S), rồi sau đó được hoàn thiện để sinh ra các loại rARN là 28S, 18S và 5,8S (*Hình 1.22b*). Riêng loại rARN 5S được phiên mã từ một gen độc lập nhờ enzym ARN polymerase III để hình thành một bản

phiên mã ngắn gồm 121 base và không phải trải qua giai đoạn hoàn thiện. Các phân tử tiền-rARN của sinh vật nhân chuẩn được hoàn thiện qua các bước tương tự như ở sinh vật nhân sơ: sau phiên mã, phân tử tiền-rARN cuộn xoắn và gắn kết với các protein cấu trúc ribosom. Sau đó, một số base bị cải biến qua quá trình methyl hoá đường ribose. Phản ứng này do các **protein nhân kích thước nhỏ** (ký hiệu là *snRNP* hoặc *snURP*) xúc tác. Giai đoạn tiếp theo là các enzym ribonuclease cắt các phân tử tiền-rARN thành các phân tử rARN 18S, 28S và 5,8S hoàn chỉnh. Vết cắt đầu tiên trên phân tử tiền-rARN 45S là ở vùng được gọi là **vùng đệm phiên mã ngoài ETS** (*e*xternal *t*ranscribed *s*pacer). Sau đó là vết cắt ở vùng được gọi là **vùng đệm phiên mã trong ITS** (*i*nternal *t*ranscribed *s*pacer) để sinh ra các phân tử tiền thân của các rARN là rARN 20S và 32S. Vết cắt tiếp theo sẽ sinh ra các phân tử rARN 28S, 18S và 5,8S hoàn chỉnh. Bước hoàn thiện cuối cùng là phân tử rARN 5,8S kết cặp bổ trợ base với phân tử rARN 28S.

I.G ARN THÔNG TIN (mARN)

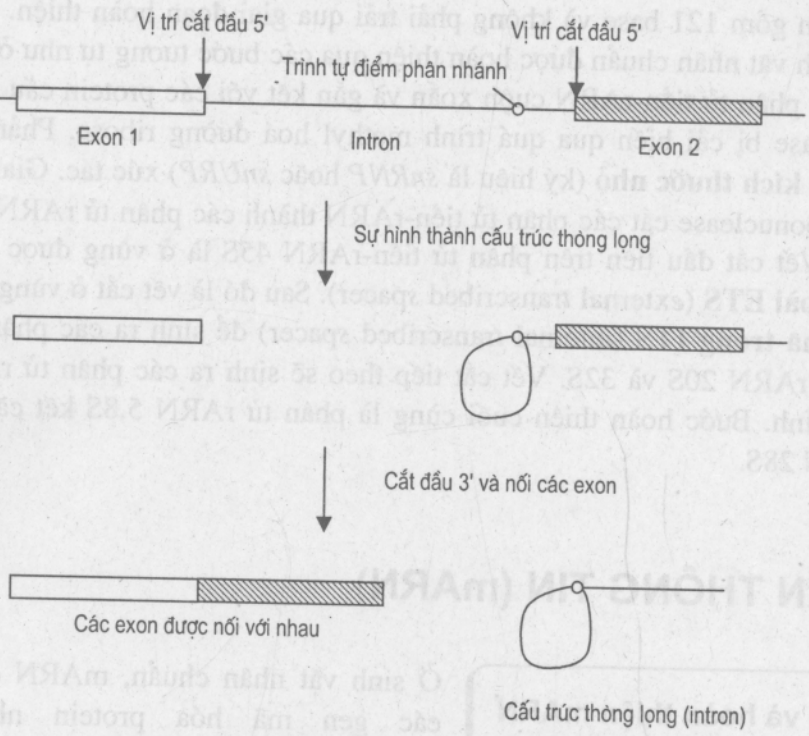
Tổng hợp và hoàn thiện mARN

Ở sinh vật nhân chuẩn, mARN được phiên mã từ các gen mã hóa protein nhờ enzym ARN polymerase II và phân tử này được dùng làm khuôn để tổng hợp protein trong quá trình dịch mã. Trình tự mã hoá của các gen sinh vật nhân chuẩn không liên tục mà được phân bố thành một loạt các exon ngăn cách nhau bởi các intron không mã hóa. Đầu tiên, bằng cách phiên mã cả exon và intron, phân tử mARN được tổng hợp dưới dạng một phân tử tiền thân, gọi tắt là tiền-mARN. Trước khi trở thành khuôn để tổng hợp protein, tiền-mARN phải trải qua quá trình hoàn thiện để trở thành mARN hoàn chỉnh. Các trình tự intron không mang mã được loại bỏ nhờ quá trình cắt bỏ làm cho trình tự mã hoá trở nên liên tục và đảm bảo cho mARN là khuôn mẫu chính xác để tổng hợp phân tử protein tương ứng. Thêm vào đó, đầu 5' của mARN được sửa đổi bằng cách thêm vào một nucleotid bị cải biến được gọi là sự lắp mũ và đầu 3' cũng được sửa đổi bằng cách thêm vào một cái đuôi chứa polyA (có thể chứa tới 250 adenin) được gọi là sự đa adenin hoá. Các mARN được phiên mã nhờ enzym ARN polymerase II có mặt trong nhân tế bào và tập hợp thành một nhóm các phân tử có độ dài khác nhau (phản ánh sự biến động về kích thước của các gen) vào những thời điểm khác nhau của quá trình hoàn thiện mARN. Nhóm các phân tử mARN như vậy được gọi là các ARN dị nguyên của nhân, viết tắt là hnARN (*h*eterogeneuos *n*uclear RNA).

Ở sinh vật nhân sơ, các phân tử mARN không phải trải qua quá trình hoàn thiện như ở sinh vật nhân chuẩn. Thậm chí quá trình dịch mã có thể bắt đầu trước khi quá trình phiên mã kết thúc. Các gen ở sinh vật nhân sơ thường không chứa các intron và vì thế không cần quá trình cắt bỏ.

Cắt bỏ các intron

Quá trình này xảy ra trong nhân nhằm cắt bỏ các trình tự intron không mã hóa khỏi phân tử tiền-mARN để hình thành nên phân tử mARN hoàn chỉnh chỉ chứa các trình tự mã hoá liên tục tương ứng với các exon. Sau đó, phân tử mARN hoàn chỉnh được chuyển ra tế bào chất để làm khuôn tổng hợp protein.

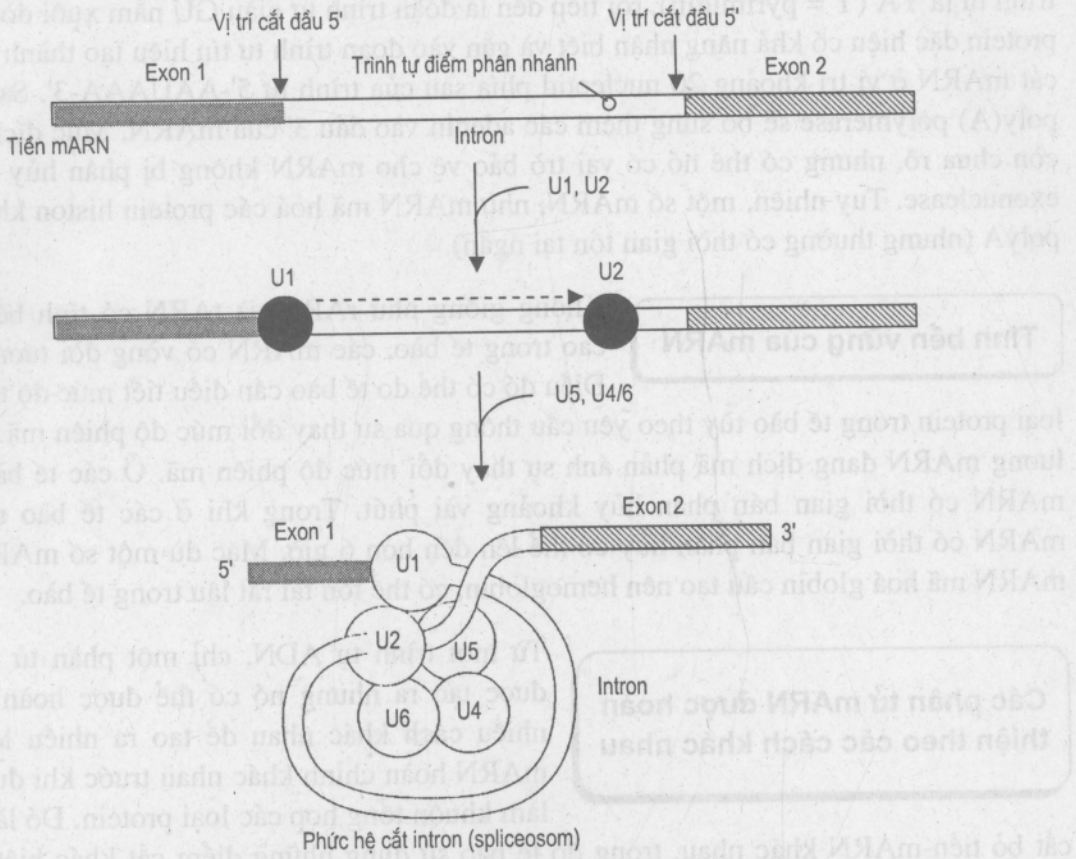


Hình 1.23. **Quá trình cắt bỏ intron của phân tử mARN tiền thân ở sinh vật nhân chuẩn**

Quá trình cắt bỏ phụ thuộc vào đoạn trình tự tín hiệu trên phân tử tiền-mARN. Ở hầu hết các gen, hai nucleotid đầu tiên ở đầu 5' của intron là GT và hai nucleotid cuối cùng ở đầu 3' là AG. Chúng là một phần của đoạn trình tự tín hiệu dài nằm ở phía đầu 3' và 5'của các intron. Đoạn trình tự tín hiệu đầy đủ ở đầu 5' là: 5'-AGGTAAGT-3' và ở đầu 3' là 5'-YYYYYYNCAG-3' (Y = pyrimidin, N = nucleotid bất kỳ). Ngoài ra, ở động vật có xương sống còn có đoạn trình tự 5'-CURAY-3' (R = purin) được gọi là đoạn trình tự điểm phân nhánh nằm ở cách khoảng 10 - 40 base phía trước (ngược dòng) trình tự tín hiệu intron ở đầu 3'. Ở nấm men có một trình tự đặc biệt ở các intron là 5'-UACUAAC-3'. Quá trình cắt bỏ diễn ra theo hai giai đoạn (*Hình 1.23*). Trong giai đoạn một, nhóm 2'-OH của adenin ở trình tự điểm phân nhánh tương tác với liên kết phosphodieste ở đầu 5' của G thuộc trình tự GT (điểm cắt đầu 5'). Trước tiên liên kết phosphodieste bị phá vỡ, giải phóng đầu 5'-G của intron và đầu này được gắn với trình tự điểm phân nhánh. Lúc này, intron hình thành một cấu trúc thòng lọng. Trong giai đoạn hai, đầu 3' của intron bị cắt sau G của trình tự AG (điểm cắt đầu 3'). Intron được giải phóng ra và hai exon được nối lại với nhau.

Quá trình cắt bỏ nhờ sự xúc tác của một nhóm các phân tử gọi là các ribonucleoprotein nhân kích thước nhỏ (snRNP). Các phân tử này được cấu tạo từ các phân tử ARN kích thước nhỏ, giàu uracil nên được gọi là U-ARN hay ARN nhân kích thước nhỏ (snARN) kết hợp với protein. Có nhiều loại snRNP, song phổ biến nhất là U1, U2, U4, U5 và U6; chúng xúc tác cho phản ứng cắt bỏ. Mỗi snRNP gồm một phân tử U-ARN, trừ U4 và U6 tồn tại dưới dạng kết cặp tương đồng của các base trong cùng một phân tử snRNP. Thành phần ARN của snRNP tương tác với đoạn trình tự tín hiệu cắt bỏ bằng cách kết cặp theo nguyên tắc bổ trợ của các base. U1 snRNP gắn với điểm cắt đầu 5'; còn U2 snRNP gắn với trình tự điểm phân nhánh. Sau đó, U5 và U4/U6 tạo thành một phức hợp với U1 và U2 làm cho intron có cấu trúc

dạng thòng lọng và các exon tiến đến gần nhau. Tổ hợp của phân tử tiền-mARN và các snRNP được gọi là phức hệ cắt intron (spliceosom) có vai trò cuộn xoắn phân tử tiền-mARN theo hình dạng cần thiết để có thể tiến hành quá trình cắt bỏ (*Hình 1.24*). Spliceosom còn có vai trò xúc tác cho phản ứng cắt intron và nối các exon. Hoạt động chức năng của spliceosom nói chung và vai trò hoạt tính enzym của từng thành phần còn chưa được hiểu biết đầy đủ.



Hình 1.24. **Sự hình thành phức hệ cắt intron (spliceosom)**

Cho dù hầu hết các intron được cắt bỏ bởi spliceosom, song cũng tồn tại một số cơ chế cắt khác nữa. Chẳng hạn, các intron trong các gen rARN ở một số sinh vật đa bào có thể hình thành dạng cấu trúc ba chiều bằng cách kết cặp các base bổ trợ. Sau đó, phức hệ này hoạt động như một enzym cắt (có bản chất là ARN và được gọi là các ribozym) tự thực hiện phản ứng cắt của mình. Các intron cũng xuất hiện ở các gen mã hóa tARN và được cắt bỏ nhờ các ribonuclease trong quá trình hoàn thiện tARN.

Lắp mũ

Đầu 5' của phân tử mARN ở sinh vật nhân chuẩn được sửa đổi bằng cách gắn thêm một nucleotid bị cải biến là 7-methylguanosin (7-mG); quá trình đó được gọi là sự lắp mũ. Mũ 7-mG được gắn nhờ enzym guanyltransferase nối GTP với nucleotid đầu tiên của mARN bằng liên kết triphotphat 5'→ 5' khác thường. Sau đó enzym methyl transferase sẽ gắn thêm nhóm -CH_3 vào nitơ số 7 của vòng guanin; đồng thời thường gắn thêm cả vào nhóm 2'-OH của đường ribose của hai nucleotid kế tiếp. Việc tạo mũ giúp bảo vệ đầu 5' của mARN không bị phân hủy bởi exonuclease trong tế bào chất, đồng thời làm tín hiệu cho ribosom nhận biết điểm bắt đầu của phân tử mARN.

Gắn đuôi poly(A)

Đầu 3' của phân tử tiền-mARN của hầu hết các sinh vật nhân chuẩn được sửa đổi bằng cách thêm vào một đoạn trình tự polyA (còn được gọi là đuôi polyA) có thể dài tới 250 base adenin. Sự sửa đổi này được gọi là đa adenin hóa và cần có một trình tự tín hiệu trên phân tử tiền-mARN. Đó là trình tự 5'-AAUAAA-3' nằm gần đầu 3' của phân tử tiền-mARN. Khoảng 11 - 20 base tiếp theo có trình tự là YA (Y = pyrimidin), rồi tiếp đến là đoạn trình tự giàu GU nằm xuôi dòng. Có nhiều protein đặc hiệu có khả năng nhận biết và gắn vào đoạn trình tự tín hiệu tạo thành một phức hệ cắt mARN ở vị trí khoảng 20 nucleotid phía sau của trình tự 5'-AAUAAA-3'. Sau đó, enzym poly(A) polymerase sẽ bổ sung thêm các adenin vào đầu 3' của mARN. Mục đích tạo đuôi A còn chưa rõ, nhưng có thể nó có vai trò bảo vệ cho mARN không bị phân hủy ở đầu 3' bởi exonuclease. Tuy nhiên, một số mARN, như mARN mã hoá các protein histon không có đuôi polyA (nhưng thường có thời gian tồn tại ngắn).

Tính bền vững của mARN

Không giống như rARN và tARN có tính bền vững khá cao trong tế bào, các mARN có vòng đời tương đối ngắn. Điều đó có thể do tế bào cần điều tiết mức độ tổng hợp các loại protein trong tế bào tùy theo yêu cầu thông qua sự thay đổi mức độ phiên mã. Sự thay đổi lượng mARN đang dịch mã phản ánh sự thay đổi mức độ phiên mã. Ở các tế bào vi khuẩn, mARN có thời gian bán phân hủy khoảng vài phút. Trong khi ở các tế bào nhân chuẩn, mARN có thời gian bán phân hủy có thể lên đến hơn 6 giờ. Mặc dù một số mARN, như các mARN mã hoá globin cấu tạo nên hemoglobin, có thể tồn tại rất lâu trong tế bào.

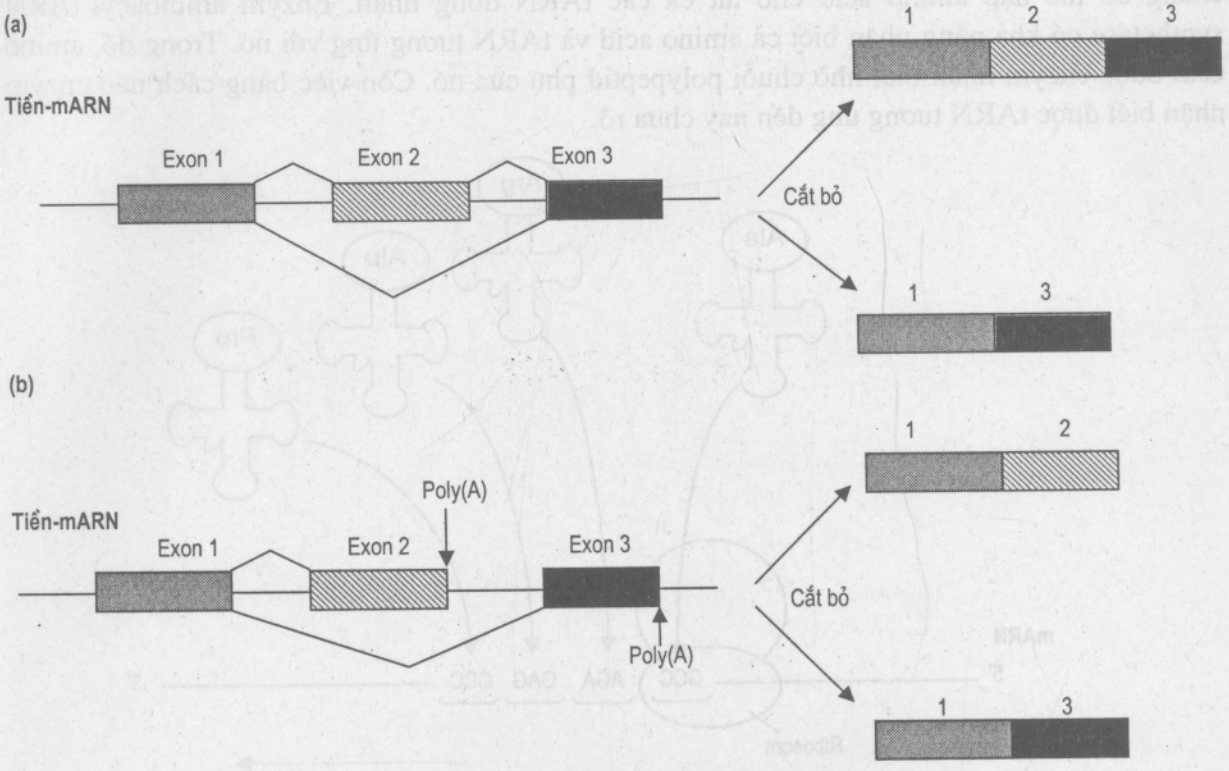
Các phân tử mARN được hoàn thiện theo các cách khác nhau

Từ một trình tự ADN, chỉ một phân tử tiền-mARN được tạo ra nhưng nó có thể được hoàn thiện bằng nhiều cách khác nhau để tạo ra nhiều loại phân tử mARN hoàn chỉnh khác nhau trước khi được sử dụng làm khuôn tổng hợp các loại protein. Đó là các cơ chế cắt bỏ tiền-mARN khác nhau, trong đó tế bào sử dụng những điểm cắt khác biệt để loại bỏ hay giữ lại các exon trong quá trình cắt bỏ. Ngoài ra, việc tồn tại các tín hiệu đa adenin hoá khác nhau trên phân tử tiền-mARN cũng có thể dẫn đến việc sinh ra các phân tử mARN có trình tự dài ngắn khác nhau ở đầu 3' (*Hình 1.25*). Ví dụ, việc sử dụng điểm đa adenin hoá nằm phía trước điểm kết thúc đoạn trình tự mã hoá có thể loại bỏ một số exon nằm sau nó và sinh ra mARN mã hóa cho một loại protein ngắn hơn.

Một phân tử tiền-mARN có thể được cắt bỏ các intron theo những cách khác nhau cùng lúc hoặc ở những giai đoạn phát triển khác nhau của một tế bào, hoặc khác nhau giữa các tế bào khác nhau. Các protein được sinh ra theo các cơ chế này thường có quan hệ với nhau, song chúng thường biểu hiện chức năng hoặc có đặc điểm riêng. Ví dụ, quá trình hoàn thiện phân tử tiền-mARN của globulin miễn dịch dẫn đến việc tổng hợp các protein có thể chứa hoặc không chứa các trình tự amino acid kỵ nước cho phép nó liên kết được vào màng tế bào. Điều này giúp tạo ra nhiều dạng globulin miễn dịch có thể liên kết với màng và các dạng để tiết ra khỏi tế bào.

Các phân tử tiền-mARN cũng còn có thể trải qua quá trình sửa đổi trình tự ARN. Trong quá trình đó, trình tự của phân tử tiền-mARN bị biến đổi bằng cách thêm vào, bớt đi hay thay thế các base. Sự sửa đổi trình tự ARN được xác định đầu tiên ở một số nguyên sinh động vật

ký sinh. Ở những loài này, người ta thấy các bản phiên mã của nhiều gen ty thể bị sửa đổi bằng cách được bổ sung thêm các gốc uracil. Quá trình này cũng gặp ở động vật có xương sống, nhưng mức độ sửa đổi ít hơn nhiều. Ở người, phân tử tiền-mARN của gen apolipoprotein B bị sửa đổi ở tế bào ruột non bằng cách thay thế base C bằng U để tạo nên một bộ ba kết thúc, dẫn đến việc tổng hợp một phân tử protein ngắn hơn. Trong khi đó ở tế bào gan, nơi trình tự ARN không bị sửa đổi, protein đó có độ dài đầy đủ.



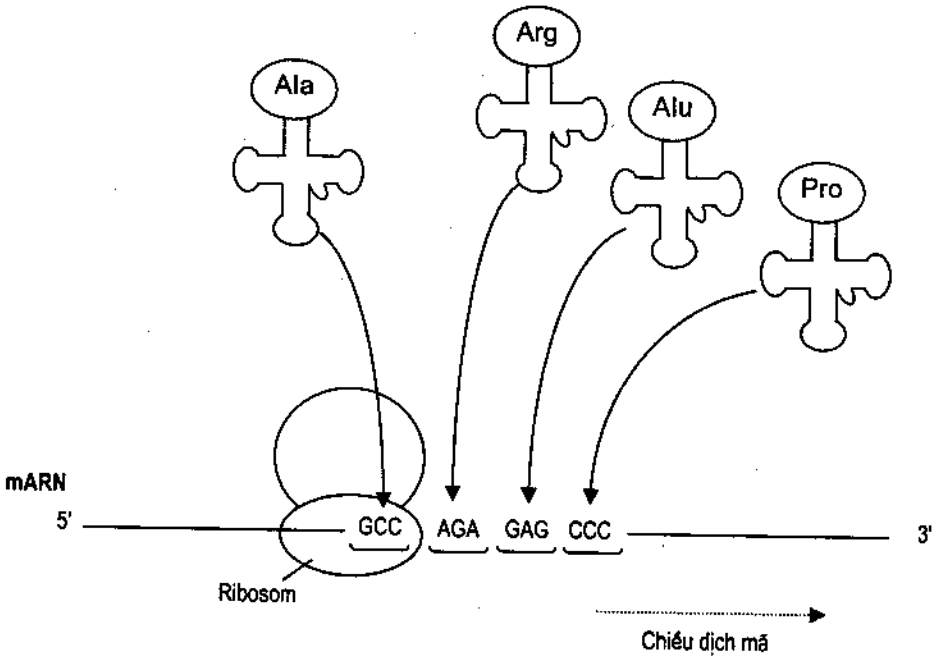
Hình 1.25. **Một số cơ chế cắt bỏ khác. (a) Bỏ qua exon, (b) Sử dụng các vị trí đa adenin hoá khác nhau**

I.H DỊCH MÃ

Vai trò của tARN

Dịch mã là quá trình tế bào tổng hợp protein. Trong quá trình này, thông tin mã hóa trên phân tử mARN được sử dụng để xác định trình tự của các amino acid trên phân tử protein. ARN vận chuyển (tARN) đóng vai trò chính trong quá trình này bởi chúng mang các amino acid đến ribosom tuân thủ theo trật tự của các nucleotid trên phân tử mARN (*Hình 1.26*). Mỗi tế bào thường có khoảng 31 - 40 loại phân tử tARN. Trong đó, mỗi phân tử tARN chỉ có thể liên kết và vận chuyển một loại amino acid duy nhất trong số 20 loại amino acid. Như vậy, để vận chuyển một loại amino acid có thể có nhiều hơn một loại tARN. Các loại phân tử tARN khác nhau cùng vận chuyển một loại amino acid được gọi là các **tARN đồng nhận**. Trước khi bắt đầu quá trình dịch mã, các amino acid liên kết đồng hoá trị với các phân tử tARN tương ứng của chúng. Sau đó, các tARN nhận biết ra các codon tương ứng trên phân tử mARN và gắn vào

phân tử này. Sự gắn kết của các amino acid với tARN thông quá phản ứng **aminoacyl hoá** (hay còn gọi là nạp amino acid). Amino acid gắn vào cuối thùy nhận amino acid của tARN nơi luôn có trình tự base là 5'-CCA-3'. Một mối liên kết đồng hóa trị hình thành giữa nhóm carboxyl của amino acid với với 3'-OH của base adenin nằm ở đầu tận cùng của thuỳ nhận amino acid. Quá trình nạp amino acid do enzym **aminoacyl-tARN synthetase** xúc tác, sử dụng năng lượng từ sự thuỷ phân ATP. Có các enzym riêng biệt cho từng loại amino acid và chúng có thể nạp amino acid cho tất cả các tARN đồng nhận. Enzym aminoacyl-tARN synthetase có khả năng nhận biết cả amino acid và tARN tương ứng với nó. Trong đó, amino acid được enzym nhận biết nhờ chuỗi polypeptid phụ của nó. Còn việc bằng cách nào enzym nhận biết được tARN tương ứng đến nay chưa rõ.



Hình 1.26. **Vai trò của tARN trong quá trình dịch mã**

Nhận biết codon

Khi đã gắn với amino acid, tARN nhận biết codon quy định amino acid đó trên phân tử mARN và đặt amino acid vào đúng vị trí của nó trên chuỗi polypeptid. Điều đó đảm bảo trình tự amino acid trên chuỗi polypeptid do phân tử mARN mã hoá được dịch mã một cách chính xác. Việc nhận biết codon diễn ra ở thùy đối mã của tARN nhờ ba nucleotid được gọi là **bộ ba đối mã** hay **anticodon** liên kết với bộ ba mã hóa bằng cách ghép đôi base theo nguyên tắc bổ trợ. Bốn loại base của ADN có thể tạo thành 64 bộ ba khác nhau. Trong đó có ba bộ ba làm tín hiệu kết thúc (là UAA, UGA và UAG), còn 61 bộ ba còn lại mã hoá cho 20 amino acid của protein. Kết quả là, hầu hết các amino acid được mã hoá bởi nhiều hơn một bộ ba. Đặc điểm này còn được gọi là **tính suy biến của mã di truyền.** Để giải quyết vấn đề này, mỗi phân tử tARN có thể nhận biết được nhiều hơn một bộ ba trên phân tử mARN quy định amino acid của nó. Hiện tượng này diễn ra được là vì bộ ba đối mã có khả năng gắn với các base khác ở vị trí thứ ba của bộ ba mã hoá. Tính chất này được gọi là **tính thoái hóa của base thứ ba** hay **tính linh hoạt** của mã bộ ba. Sở dĩ sự gắn kết giữa codon và anticodon có thể cho phép có sự biến động

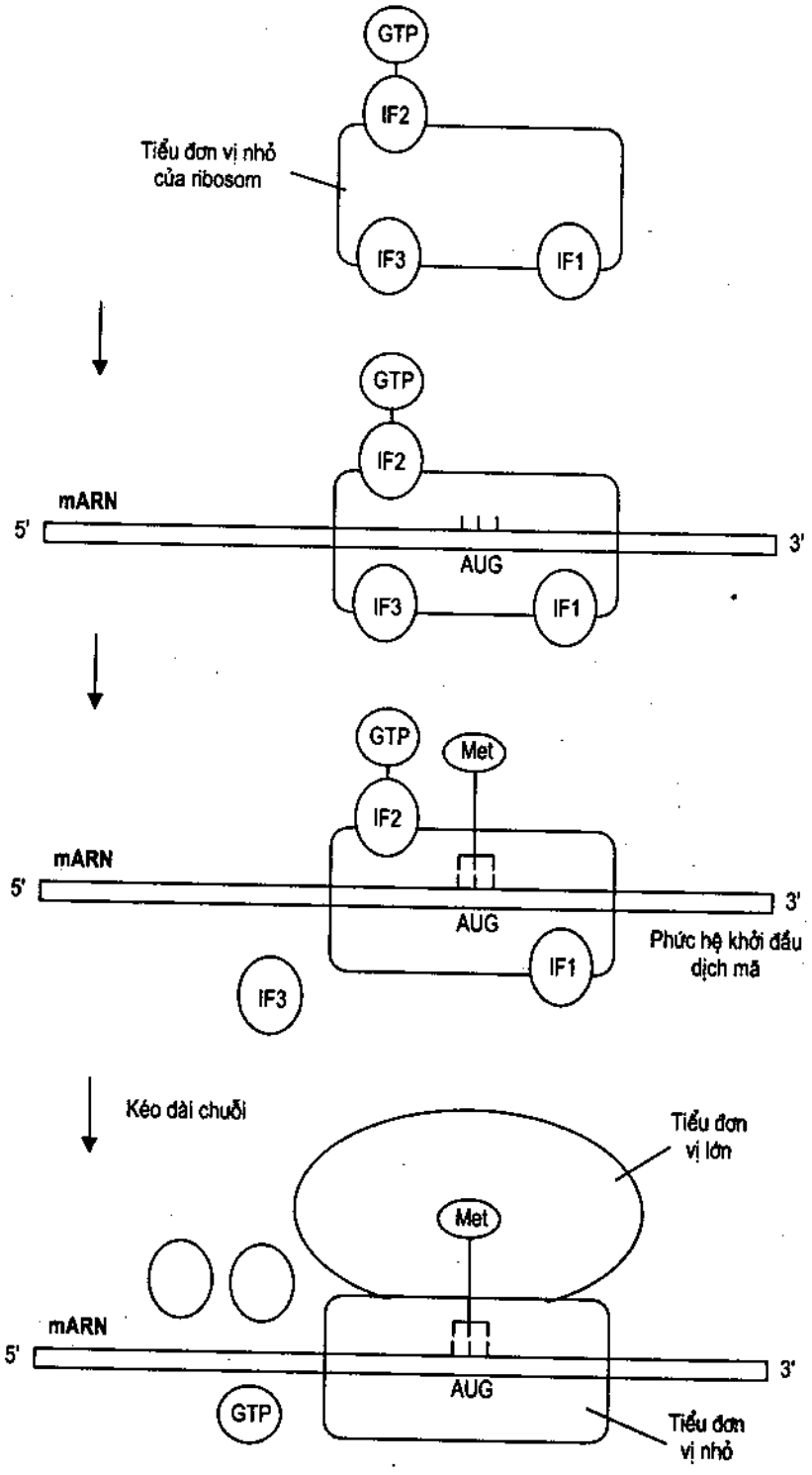
ở base thứ ba vì thùy đối mã không thẳng ở vị trí này và khi anticodon gắn kết với codon không có sự hình thành nên chuỗi xoắn kép ARN hoàn chỉnh. Điều đó tạo điều kiện cho sự hình thành một số kiểu liên kết đôi base không chuẩn. Ví dụ, G có thể kết với U và với C ở vị trí linh hoạt; hoặc inosin (một dạng khử amin của guanin, đôi khi có mặt ở vị trí linh hoạt của thùy đối mã) có thể liên kết không chỉ với C mà cả với A và U. Tính linh hoạt cho phép một phân tử tARN có thể giải mã nhiều hơn một codon, nhờ vậy làm giảm số loại tARN mà tế bào cần. Tuy vậy, các nguyên tắc cơ bản của mã di truyền không bị xâm phạm và các chuỗi polypeptid đều được tổng hợp chính xác theo trình tự của mARN. Một hệ quả quan trọng của tính linh hoạt là nó góp phần làm giảm thiểu hiệu ứng của đột biến.

Dịch mã

Quá trình dịch mã ở sinh vật nhân sơ và sinh vật nhân chuẩn tương tự nhau và đều được chia làm ba giai đoạn: **khởi đầu, kéo dài** và **kết thúc**. Sự khởi đầu dịch mã bao gồm việc ribosom gắn vào phân tử mARN. Quá trình kéo dài dịch mã bao gồm việc bổ sung các amino acid vào chuỗi polypeptid và quá trình kết thúc dịch mã liên quan đến việc giải phóng chuỗi polypeptid vừa được tổng hợp. Có nhiều loại **protein trợ giúp** tham gia vào các giai đoạn của quá trình dịch mã. Trong quá trình dịch mã, tế bào sử dụng năng lượng thu được do quá trình thuỷ phân guanosin triphosphat (GTP) và adenosin triphosphat (ATP). GTP được dùng cho sự vận động của các protein và liên kết các yếu tố dịch mã. ATP được dùng để nạp amino acid vào phân tử tARN và để gỡ bỏ cấu trúc bậc hai của mARN. Người ta xác định có đến 90% lượng ATP mà vi khuẩn sinh ra được dùng để đáp ứng yêu cầu của các hoạt động dịch mã.

Khởi đầu dịch mã

Khi không tham gia dịch mã, các tiểu đơn vị lớn và nhỏ của ribosom tồn tại độc lập. Trong giai đoạn đầu tiên của quá trình dịch mã, tiểu đơn vị nhỏ của ribosom gắn kết vào phân tử mARN (*Hình 1.27*). Sự dịch mã thường bắt đầu từ trình tự AUG (đôi khi ở vi khuẩn là GUG hoặc UUG). Mã bộ ba này đồng thời mã hóa cho methionin và được gọi là bộ ba khởi đầu dịch mã. Tiểu đơn vị nhỏ của ribosom thường bắt đầu gắn vào phân tử mARN ở một vị trí đặc hiệu nằm phía trước (ngược dòng) bộ ba AUG. Ở sinh vật nhân sơ, trình tự đó là Shine-Dalgarno (5'-AGGAGGU-3') nằm gần điểm bắt đầu của phân tử mARN. Khi gắn vào phân tử mARN, tiểu đơn vị nhỏ của ribosom di chuyển dọc phân tử mARN theo chiều 5' → 3' cho đến khi gặp bộ ba AUG (thường xuôi dòng khoảng 10 nucleotid). Ở sinh vật nhân chuẩn, tiểu đơn vị nhỏ có thể nhận biết được mũ ở đầu 5' của phân tử mARN. Sau đó nó dịch chuyển xuôi dòng cho đến khi bắt gặp bộ ba AUG đầu tiên, dù rằng đôi khi cũng có thể nhận nhầm những bộ ba AUG khác. Phân tử tARN mang methionin gắn vào bộ ba AUG đã được tiểu đơn vị nhỏ của ribosom định vị. Ở vi khuẩn, methionin bị cải biến bằng cách thêm nhóm formyl (-CHO) vào một trong số những nguyên tử hydro của nhóm amin (khi đó tARN được ký hiệu là $tARN^{fMet}$). Sự kiện đó nhằm khóa nhóm amin để ngăn cản sự hình thành liên kết peptid về phía đầu N của chuỗi polypeptid và đảm bảo cho quá trình polyme hoá chuỗi polypeptid chỉ theo hướng từ đầu N tới đầu C. Khi kết hợp với mARN, tiểu đơn vị nhỏ của ribosom và $tARN^{fMet}$ tạo thành phức hệ khởi đầu dịch mã. Có hai loại phân tử tARN có thể nhận biết bộ ba AUG và mang methionin. Trong đó, một loại được dùng để khởi đầu dịch mã, còn loại tARN kia nhận biết các bộ ba AUG khác (mã hóa cho methionin) bên trong chuỗi. Chỉ có tARN khởi đầu dịch mã mới gắn được vào phức hệ khởi đầu dịch mã.



Bước 1. Tiểu đơn vị nhỏ của ribosom gắn vào các yếu tố IF1, IF3 và phức hợp IF2 + GTP.

Bước 2. Tiểu đơn vị nhỏ gắn vào phân tử mARN và trượt đến vị trí codon khởi đầu dịch mã AUG

Bước 3. tARN mang methionin khởi đầu dịch mã gắn vào codon AUG. IF3 được giải phóng

Bước 4. Tiểu đơn vị lớn gắn vào ribosom. IF1, IF2 được giải phóng. GTP bị thủy phân

Hình 1.27. **Sự khởi đầu dịch mã ở *E. coli***

Để khởi đầu dịch mã, tế bào còn cần nhiều loại protein trợ giúp, gọi là các yếu tố khởi đầu dịch mã. Ở vi khuẩn có ba yếu tố là IF1, IF2 và IF3. Sự khởi đầu dịch mã bắt đầu bằng việc IF1 và IF3 gắn vào tiểu đơn vị nhỏ của ribosom. Điều này giúp ngăn cản tiểu đơn vị nhỏ liên kết với tiểu đơn vị lớn trước khi tiểu đơn vị nhỏ gắn được vào phân tử mARN. Sau đó,

phức hợp IF2 và GTP gắn vào tiểu đơn vị nhỏ và thúc đẩy sự liên kết của phân tử tARN khởi đầu dịch mã vào tiểu đơn vị nhỏ. Tiếp theo sau, tiểu đơn vị nhỏ gắn vào phân tử mARN và định vị bộ ba khởi đầu dịch mã AUG. Lúc này, phân tử tARN mang methionin khởi đầu dịch mã có thể gắn được vào phức hệ khởi đầu dịch mã, đồng thời giải phóng ra yếu tố IF3. Sự kiện này đánh dấu việc kết thúc quá trình khởi đầu dịch mã. Quá trình kéo dài dịch mã được bắt đầu bằng việc liên kết của tiểu đơn vị lớn của ribosom với phức hệ khởi đầu dịch mã, đồng thời tạo ra một ribosom hoàn chỉnh. Song song với sự kiện này là việc giải phóng các yếu tố IF1, IF2 và thủy phân GTP (*Hình 1.27*).

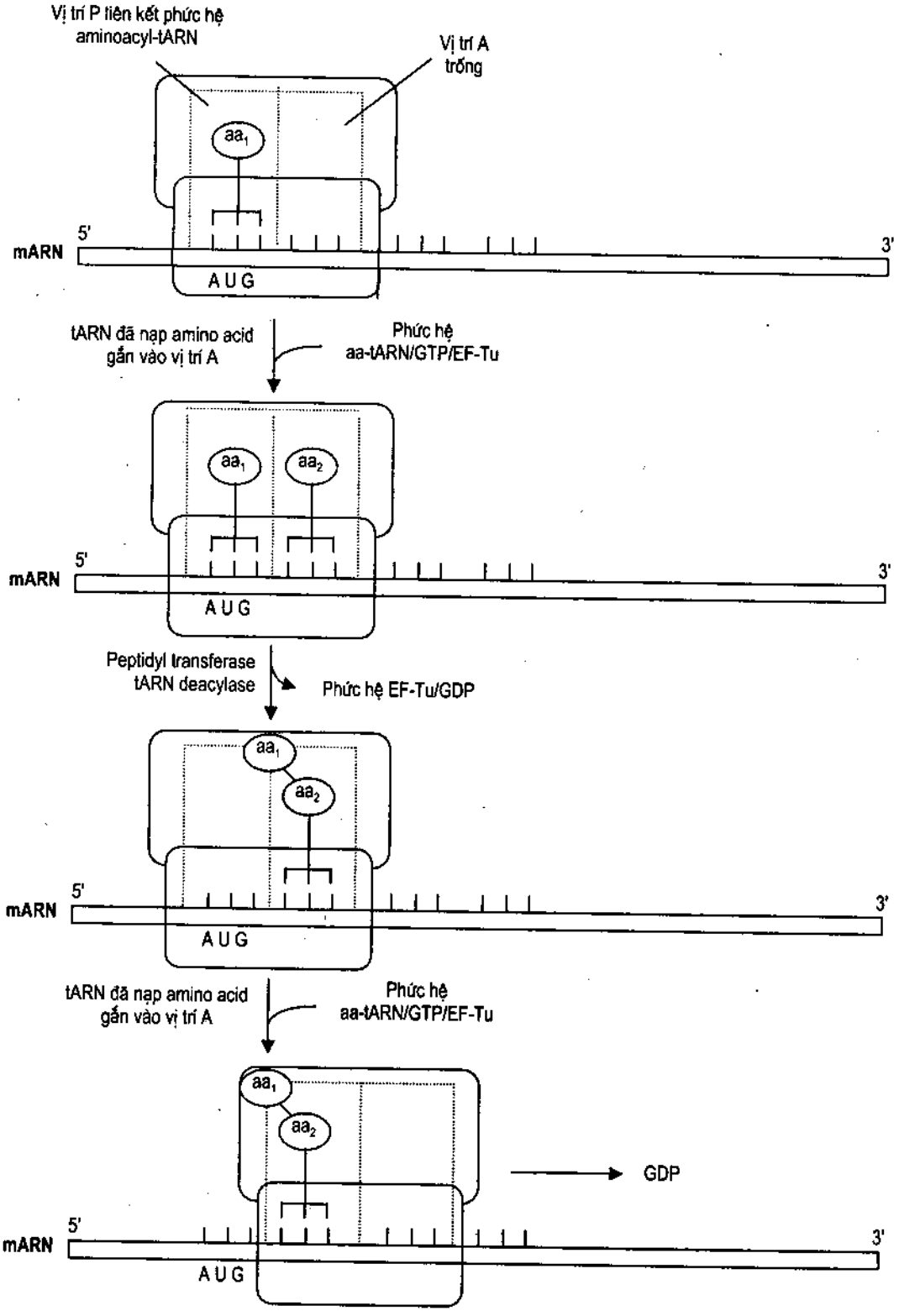
Sự khởi đầu dịch mã ở sinh vật nhân chuẩn gần giống như ở sinh vật nhân sơ. Tuy nhiên, bộ ba khởi đầu dịch mã hầu như luôn luôn là AUG và amino acid đầu tiên (methionin) không bị formyl hoá. Có nhiều (ít nhất là 9) yếu tố tham gia vào quá trình khởi đầu dịch mã. Trong đó, một số yếu tố có vai trò giống như các yếu tố tương ứng ở sinh vật nhân sơ. Chẳng hạn, hai yếu tố eIF2 và eIF3 ở sinh vật nhân chuẩn có vai trò giống với IF2 và IF3 ở vi khuẩn. Một số yếu tố tham gia vào việc gỡ bỏ cấu trúc bậc hai của phân tử mARN trước khi quá trình dịch mã có thể bắt đầu. Năng lượng cần cho quá trình này thu được từ sự thuỷ phân ATP.

Kéo dài dịch mã

Ngay sau khi phức hệ khởi đầu dịch mã hình thành, tiểu đơn vị lớn của ribosom sẽ gắn vào nó. Lúc này ribosom hoàn chỉnh có hai vị trí gắn cho các phân tử tARN (*Hình 1.28*). Vị trí thứ nhất là vị trí P (vị trí peptidyl) do $tARN_i^{Met}$ liên kết bổ trợ với bộ ba AUG chiếm giữ. Vị trí thứ hai là vị trí A (vị trí aminoacyl) tương ứng với bộ ba thứ hai trên phân tử mARN. Quá trình kéo dài dịch mã bắt đầu khi phân tử tARN thứ hai gắn vào vị trí A và liên kết với bộ ba thứ hai. Khi cả hai vị trí đều bị các phân tử tARN chiếm giữ thì các amino acid được sắp xếp kề nhau và lúc này mối liên kết peptid được hình thành giữa nhóm carboxyl của amino acid thứ nhất (methionin) với nhóm amin của amino acid thứ hai, nhờ hoạt động xúc tác của enzym peptidyl transferase. Enzym peptidyl transferase cùng hoạt động với một enzym khác là tARN deacylase. Enzym thứ hai này làm nhiệm vụ phá huỷ mối liên kết giữa amino acid thứ nhất và phân tử tARN mang nó sau khi mối liên kết peptid hình thành.

Tế bào cũng cần nhiều protein trợ giúp để quá trình kéo dài dịch mã có thể diễn ra. Ở sinh vật nhân sơ, có hai yếu tố kéo dài dịch mã là EF-Tu và EF-Ts. EF-Tu tạo phức với GTP và liên kết với phân tử tARN lúc này đã được nạp amino acid (gọi là aminoacyl-tARN) khi phân tử này gắn vào vị trí A trên ribosom. Sau khi aminoacyl-tARN đã gắn vào vị trí A, GTP bị thủy phân và EF-Tu liên kết với guanosin diphosphat (GDP) được giải phóng ra. Trước khi một phân tử tARN khác có thể gắn với GTP, yếu tố EF-Tu phải được tái tạo nhờ yếu tố EF-Ts. Trước tiên, EF-Ts thay thế GDP để gắn vào EF-Tu, rồi sau đó phân tử GTP mới sẽ thay thế EF-Ts để tạo thành phức hợp với EF-Tu (*Hình 1.29*). Ở sinh vật nhân chuẩn, một protein phức tạp gọi là eEF-1 mang tARN vào vị trí A. Phản ứng này cũng đi kèm với phản ứng thủy phân GTP. Chi tiết về cách tái tạo eEF-1 còn chưa rõ, song có thể có các thành phần tương ứng với EF-Tu và EF-Ts.

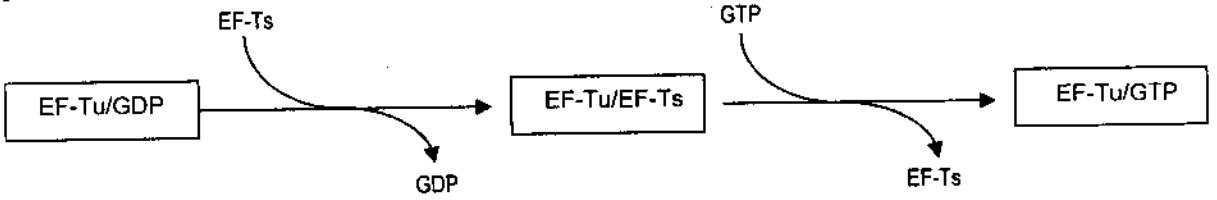
Sau khi hình thành liên kết peptid, ribosom dịch chuyển đến codon liền kề tiếp sau. Liên kết dipeptid mới hình thành đang gắn với phức hệ aminoacyl-tARN thứ hai (lúc này đã là aa_1-aa_2-tARN) dịch chuyển sang vị trí P, đồng thời đẩy tARN không mang amino acid ra ngoài và để lại vị trí A trống. Phân tử tARN mang amino acid thứ ba sẽ gắn vào vị trí A trống này,



Hình 1.28. **Quá trình kéo dài dịch mã**

và cứ như thế chu trình kéo dài dịch mã được lặp đi lặp lại. Sau mỗi lần bổ sung thêm một amino acid vào chuỗi polypeptid đang tổng hợp, ribosom lại chuyển dịch một codon tiếp theo. Ở vi khuẩn, sự dịch chuyển của ribosom do yếu tố kéo dài dịch mã là EF-G xúc tác, sau khi

yếu tố này tạo phức với GTP. Trong quá trình kéo dài dịch mã, GTP bị thủy phân để cung cấp năng lượng cho sự dịch chuyển ribosom. Sự gắn kết của EF-G và EF-Tu vào ribosom có hiệu ứng loại trừ lẫn nhau. Điều đó đảm bảo cho sự dịch chuyển của ribosom diễn ra trước khi sự kéo dài dịch mã tiếp theo bắt đầu. Ở sinh vật nhân chuẩn, yếu tố tương đương với EF-G là eEF-2. Yếu tố này cũng có khả năng gắn kết với GTP và hoạt động chức năng tương tự. Trong quá trình dịch mã, ribosom dịch chuyển dọc theo phân tử mARN và ngày càng xa dần về phía đầu 3', đồng thời để lại điểm khởi đầu dịch mã tự do trên phân tử mARN cho các phức hệ ribosom mới gắn vào. Nhờ vậy, cùng lúc có thể có nhiều ribosom cùng tiến hành dịch mã một phân tử mARN, tạo thành cấu trúc gọi là **polysom**.



Hình 1.29. **Tái sinh phức hợp EF-Tu/GTP**

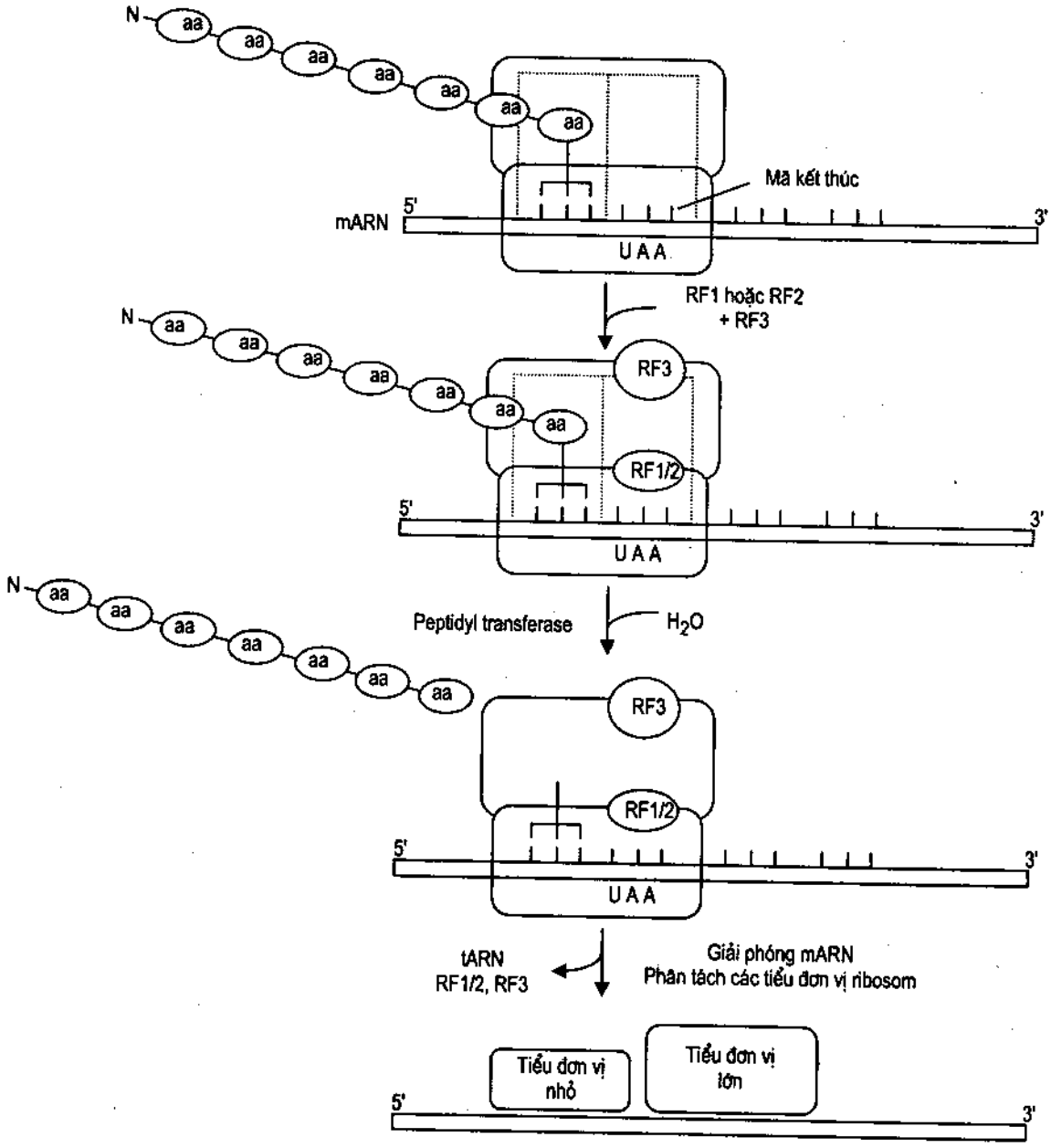
Kết thúc dịch mã

Quá trình dịch mã kết thúc khi ribosom tiếp xúc với bộ ba kết thúc ở vị trí A (*Hình 1.30*). Lúc này, không có phân tử tARN nào có thể gắn kết được với bộ ba này. Thay vào đó, các protein được gọi là các yếu tố kết thúc dịch mã (còn gọi là các yếu tố giải phóng chuỗi) sẽ xâm nhập vị trí A và góp phần giải phóng chuỗi polypeptid hoàn chỉnh ra khỏi phức hệ dịch mã. Ở *E. coli* có ba yếu tố giải phóng chuỗi, ký hiệu là RF1, RF2 và RF3. RF1 nhận biết các bộ ba kết thúc UAA và UAG, trong khi RF2 nhận biết các bộ ba UAA và UGA. Riêng yếu tố RF3 đóng vai trò phụ trợ trong quá trình này. Các yếu tố kết thúc dịch mã làm enzym peptidyl transferase xúc tác việc liên kết chuỗi polypeptid với một phân tử nước thay cho một phức hệ aminoacyl-tARN mới.

Ở sinh vật nhân chuẩn, chỉ phát hiện thấy có một protein là eRF tham gia vào quá trình kết thúc dịch mã. Yếu tố này tạo phức với GTP trước khi gắn vào ribosom. Sau đó, GTP bị thủy phân và eRF được giải phóng khỏi ribosom. Sau khi giải phóng khỏi chuỗi polypeptid, ribosom tiếp tục giải phóng khỏi phân tử mARN và tự phân tách. Lúc này, các tiểu đơn vị ribosom lớn và nhỏ tách nhau ra có thể được dùng để tham gia vào lần dịch mã sau.

Sửa đổi sau dịch mã

Sau quá trình dịch mã, các chuỗi polypeptid mới được tổng hợp có thể được biến đổi trước khi trở thành phân tử protein có đầy đủ chức năng. Những biến đổi đó có thể là sự gắn thêm các nhóm chất thông qua liên kết đồng hoá trị, hoặc bằng các phản ứng cắt bỏ bớt một số trình tự trên chuỗi polypeptid. Trong số những sửa đổi này, người ta tìm thấy các sửa đổi liên quan đến các chuỗi phụ amino acid, hoặc sửa đổi ở phía đầu amin hoặc đầu carboxyl. Sự sửa đổi có thể liên quan đến việc bổ sung những nhóm chất nhỏ, như sự methyl hóa, phosphoryl hóa, acetyl hóa hoặc hydroxyl hoá, hay sự gắn thêm những nhóm phân tử lớn, như lipid hoặc oligosaccharid (glycosyl hoá). Một số sửa đổi protein như sự phosphoryl hoá liên quan đến sự điều chỉnh hoạt tính enzym. Việc cắt bớt chuỗi polypeptid là một kiểu sửa đổi rất phổ biến. Việc cắt bỏ này có thể liên quan đến các amino acid ở các đầu của chuỗi polypeptid, hoặc các đoạn trình tự peptid bên trong chuỗi, hoặc cắt bỏ đi trình tự peptid tín hiệu ở phía đầu amino của các phân tử protein tiết, hay cắt các phân tử đa protein thành các đoạn peptid có kích thước nhỏ. Đôi khi, việc cắt protein còn là cơ chế để hoạt hóa các phân tử protein tiền thân không hoạt động.



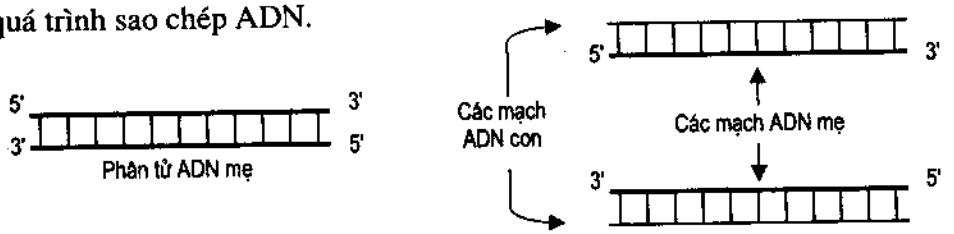
Hình 1.30. **Sự kết thúc dịch mã ở *E. coli***

I.I SAO CHÉP ADN

Sao chép ADN

Đây là quá trình tế bào sao chép lại vật chất di truyền (ADN) của nó. Quá trình này (còn gọi là quá trình tái bản ADN) là cần thiết để đảm bảo thông tin di truyền có trong tế bào được truyền cho các tế bào con sau quá trình phân bào. ADN được sao chép nhờ hoạt động của các enzym ADN polymerase. Các enzym này tác động lên một mạch đơn của phân tử ADN để tổng hợp nên một mạch mới có trình tự nucleotid bổ trợ với mạch làm khuôn. Chiều tổng hợp mạch mới luôn là 5' → 3'.

Sự sao chép ADN diễn ra theo nguyên tắc bán bảo toàn, nghĩa là mỗi bản sao phân tử ADN con có một mạch bắt nguồn từ phân tử ADN bố, mẹ và một mạch được tổng hợp mới. Cơ chế sao chép ADN rất giống nhau ở hầu hết các loài sinh vật. Điểm khác biệt chủ yếu chỉ liên quan đến thành phần các loại enzym và protein tham gia vào quá trình sao chép. Ở sinh vật nhân sơ, như *E. coli* chẳng hạn, có hai enzym là ADN polymerase I và III chịu trách nhiệm tổng hợp ADN. Ở các sinh vật nhân chuẩn có nhiều loại là ADN polymerase, chẳng hạn ADN polymerase α, β, γ, δ và ε. Quá trình sao chép ADN phải diễn ra thật chính xác vì chỉ một sai sót nhỏ cũng có thể dẫn đến việc mất đi thông tin di truyền quan trọng chỉ sau vài lần phân bào. Các enzym ADN polymerase ngoài chức năng tổng hợp ADN, còn có khả năng kiểm tra lại trình tự các base ở mạch mới được tổng hợp để đảm bảo sự sao chép chính xác. Chức năng này có được là nhờ hoạt tính exonuclease ngược (3' → 5') của các enzym. Hoạt tính này cho phép chúng có thể loại bỏ những base được lắp ráp sai trên mạch mới được tổng hợp và thay thế vào những vị trí đó những base đúng. Chức năng này gọi là chức năng đọc sửa của các enzym ADN polymerase. Ước tính trung bình, chỉ khoảng $1/(5x10^9)$ nucleotid được lắp không đúng sau quá trình sao chép ADN.



Hình 1.31. **Sao chép theo nguyên tắc bán bảo toàn**

Chạc sao chép

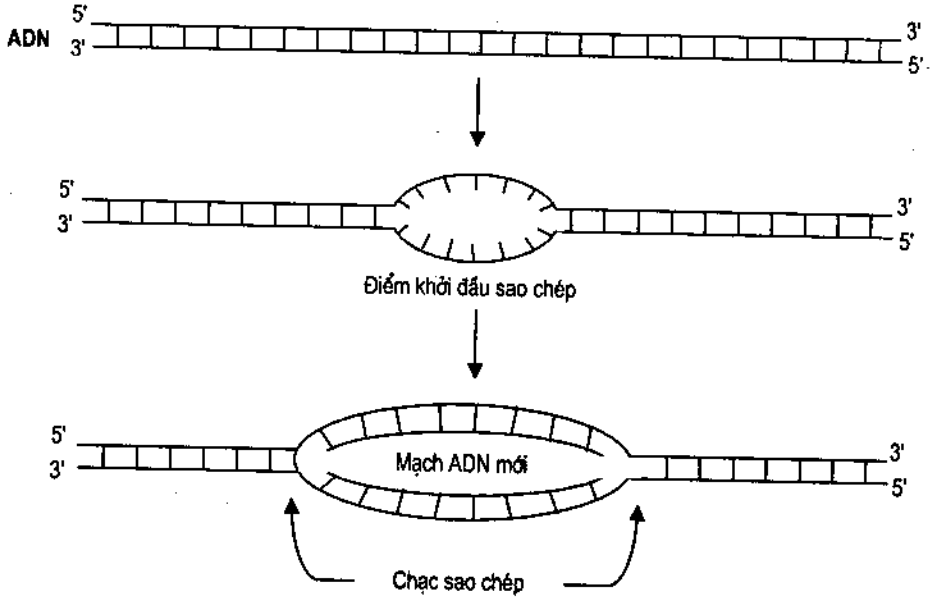
Quá trình sao chép diễn ra ở chạc sao chép. Enzym helicase tách chuỗi xoắn kép thành hai mạch đơn, và protein bám mạch đơn (SSB) giữ cho các mạch đơn tách nhau ra. ADN được tổng hợp liên tục ở mạch dẫn đầu hoặc thành từng đoạn ngắn (gọi là các đoạn Okazaki) không liên tục ở mạch ra chậm (mạch theo sau). Ở *E. coli*, ADN được tổng hợp nhờ enzym ADN polymerase III và khởi đầu bằng một đoạn mồi ARN ngắn do ARN polymerase và primase tổng hợp nên. Sau đó, trình tự mồi ARN được thay thế bằng ADN nhờ enzym ADN polymerase I. Ở sinh vật nhân chuẩn, ADN polymerase α khởi đầu quá trình tổng hợp ADN nhờ hoạt tính primase tích hợp của nó. Sợi dẫn đầu và sợi theo sau được tổng hợp tương ứng nhờ hoạt động của các enzym ADN polymerase δ và α. Enzym ADN ligase có vai trò xúc tác việc hình thành liên kết phosphodieste nối giữa các đoạn Okazaki.

Sao chép ADN ở vi khuẩn

Trong quá trình sao chép, toàn bộ phân tử ADN của tế bào được giãn xoắn từng phần để tạo ra những đoạn ADN mạch đơn cho phép các enzym ADN polymerase tiến hành quá trình sao chép. Sự giãn xoắn chuỗi xoắn kép bắt đầu từ một điểm khởi đầu sao chép và sau đó lan tỏa dần ra, thông thường theo cả hai hướng suốt dọc chiều dài phân tử ADN. Các điểm khởi đầu sao chép thường chứa đoạn trình tự giàu cặp A=T là loại liên kết yếu. Vùng trình tự mà ở đó chuỗi xoắn kép tháo xoắn và diễn ra sự tổng hợp các mạch ADN mới được gọi là chạc sao chép (*Hình 1.32*). Tại chạc sao chép có một số sự kiện xảy ra như sau:

– *Tách chuỗi xoắn kép*. Hai mạch đơn của chuỗi xoắn kép tách nhau dưới tác động của enzym helicase. Sau khi hai mạch tách ra, protein bám sợi đơn (SSB) gắn với mỗi mạch để ngăn cản việc tái tạo chuỗi xoắn.

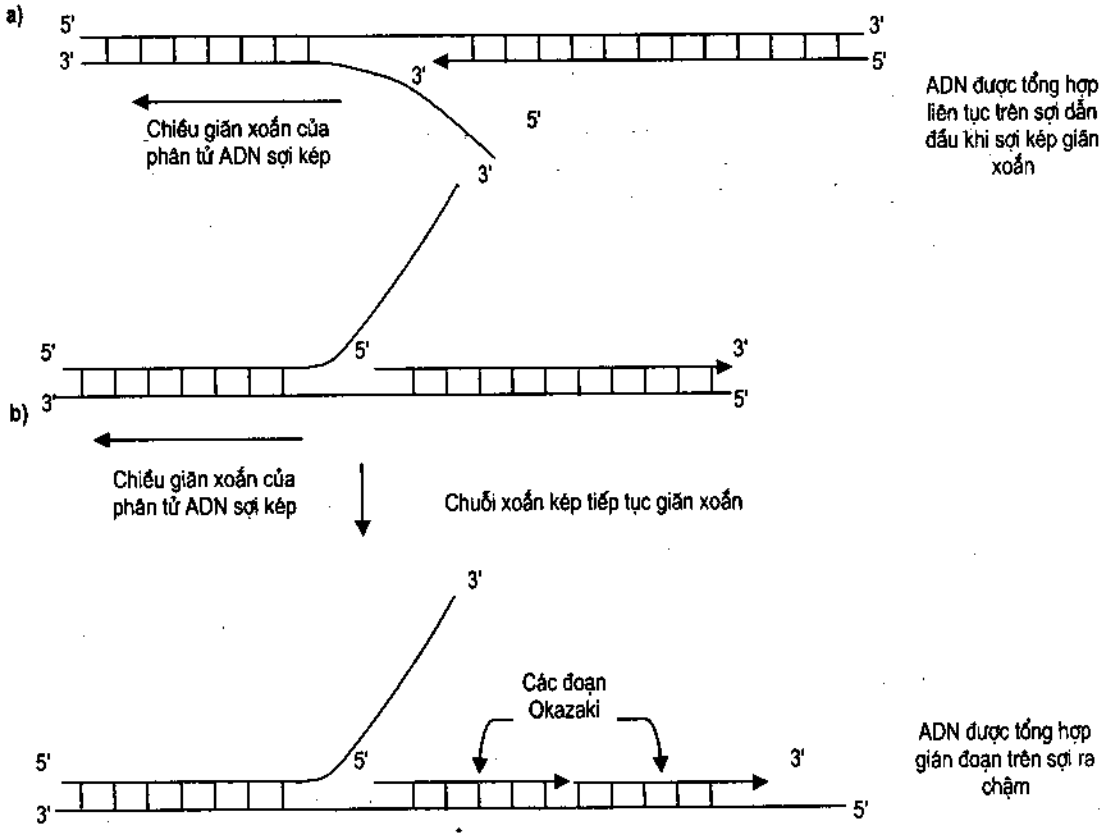
– *Tổng hợp sợi dẫn đầu và sợi theo sau*. Quá trình tổng hợp các mạch ADN được thực hiện nhờ enzym ADN polymerase chỉ diễn ra theo chiều 5' → 3'. Vì hai mạch của chuỗi xoắn kép chạy theo hai hướng ngược nhau nên cơ chế sao chép hai sợi ADN mới có khác nhau. Một sợi được gọi là sợi dẫn đầu có mạch ADN mới được sao chép đúng hướng hoạt động của enzym ADN polymerase, và vì thế nó được tổng hợp một cách liên tục (*Hình 1.33a*). Sợi kia, được gọi là sợi ra chậm (sợi theo sau) có mạch mới được tổng hợp theo chiều ngược lại và không liên tục, thay vào đó nó được tổng hợp thành nhiều đoạn nhỏ gọi là các đoạn Okazaki (*Hình 1.33b*).



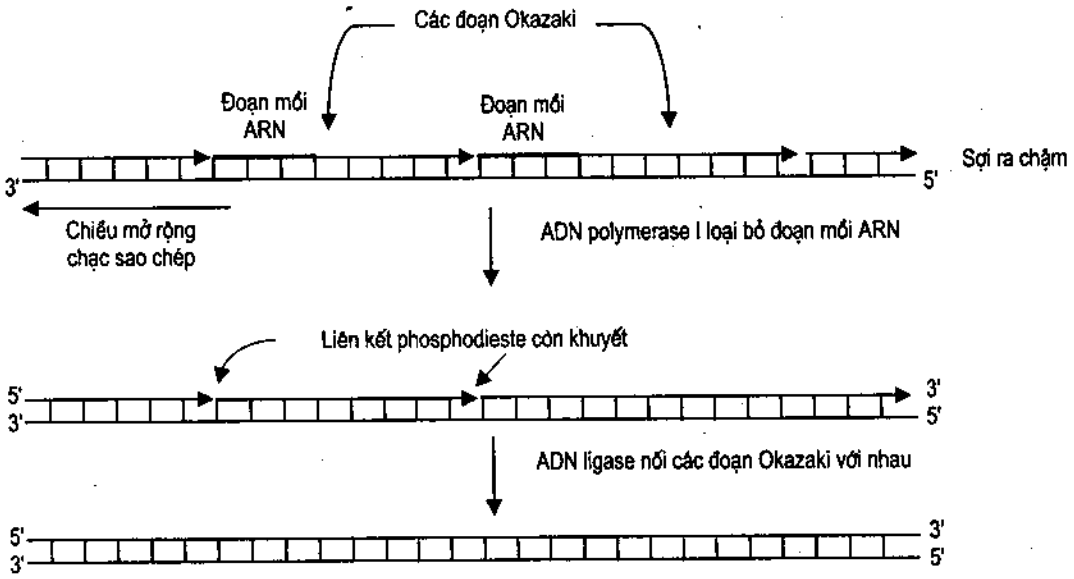
Hình 1.32. **Điểm khởi đầu sao chép và chạc sao chép**

– *Tạo đoạn mồi*. Các enzym ADN polymerase chỉ có thể bắt đầu tiến hành sao chép ADN khi đã có sẵn một đoạn sợi kép ngắn gọi là đoạn mồi. Quá trình tạo đoạn mồi do một enzym có tính chất ARN polymerase có khả năng khởi đầu sự tổng hợp trên mạch đơn ADN là primase thực hiện. Primase tổng hợp một đoạn trình tự ARN mồi ngắn dựa trên sợi ADN khuôn theo nguyên tắc bổ trợ. Ở *E. coli*, ADN polymerase III sau đó sẽ tiến hành tổng hợp ADN bắt đầu từ đoạn mồi ARN. Trên sợi ra chậm, quá trình tổng hợp sẽ dừng lại khi bắt gặp đoạn ARN mồi kế tiếp. Lúc này, enzym ADN polymerase I sẽ làm nhiệm vụ tách bỏ đoạn ARN mồi và thay thế vào nó bằng trình tự ADN tương ứng (*Hình 1.34*). Ở sinh vật nhân chuẩn, có một đặc điểm khác là enzym ADN polymerase α có hoạt tính primase tích hợp chịu trách nhiệm tổng hợp đoạn mồi này. Sau đó, hai enzym ADN polymerase α và δ chịu trách nhiệm tiếp tục tổng hợp phân tử ADN. Trong đó ADN polymerase δ tổng hợp sợi dẫn đầu, còn ADN polymerase α tổng hợp sợi theo sau. Các enzym polymerase khác có vai trò phụ trợ. ADN polymerase ε tham gia vào sửa chữa ADN và ADN polymerase γ có vai trò sao chép ADN ty thể.

– *Kết nối các đoạn Okazaki*. Bước hoàn chỉnh quá trình tổng hợp sợi theo sau là kết nối các đoạn Okazaki bằng sự hình thành liên kết photphodieste nhờ enzym ADN ligase (*Hình 1.34*).



Hình 1.33. **(a) Sao chép mạch dẫn đầu. (b) Sao chép mạch ra chậm (mạch theo sau)**

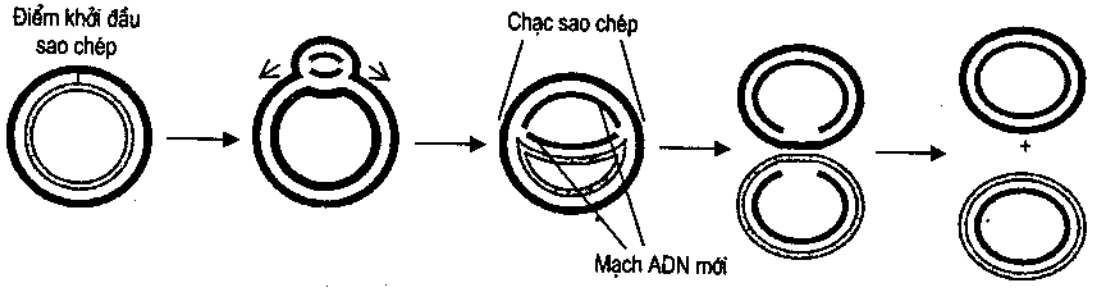


Hình 1.34. **Hoàn thành tổng hợp sợi ra chậm (sợi theo sau)**

Dù nguyên lý sao chép ADN cơ bản giống nhau ở mọi sinh vật, nhưng vẫn có những đặc điểm sao chép khác nhau phụ thuộc vào bản chất của phân tử ADN được sao chép. Các phân

tử ADN ở vi khuẩn có cấu trúc mạch vòng có đặc điểm sao chép khác với các phân tử ADN mạch thẳng ở sinh vật nhân chuẩn. Kiểu sao chép đơn giản và phổ biến nhất đối với các phân tử ADN mạch vòng là từ chỉ một điểm khởi đầu sao chép hình thành nên hai chạc sao chép di chuyển về hai hướng ngược nhau. Sự kiện đó làm cho phân tử ADN có dạng hình thêta θ (*Hình 1.35*). Cuối cùng, các chạc sao chép gặp nhau, hòa hợp lại và quá trình sao chép kết thúc.

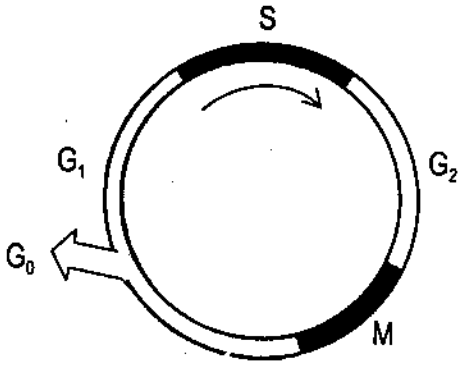
Để có thể sao chép, phân tử ADN cần được tháo xoắn. Để tháo xoắn, đoạn phân tử ADN phía trước chạc sao chép phải xoay tròn. Vì phân tử ADN mạch vòng không có các đầu tự do, nên quá trình tháo xoắn như vậy có thể gây rối phân tử ADN và ngăn cản chạc sao chép tiến về phía trước. Trở ngại này được giải quyết nhờ hai enzym topoisomerase. Trong đó, enzym topoisomerase I làm gẫy bộ khung của chuỗi polynucleotid trên một mạch của phân tử ADN cách chạc sao chép một đoạn ngắn về phía trước. Điều này cho phép một mạch ADN có thể quay tròn quanh sợi kia. Sau đó, chính enzym topoisomerase I nối lại các đầu bị đứt. Quá trình sao chép kết thúc sẽ tạo ra hai phân tử con đan xen vào nhau. Lúc này, ADN topoisomerase II bằng cách gây ra những vết đứt cách đoạn trên cả hai mạch của một phân tử ADN con để tạo điều kiện cho phân tử kia tách ra. Sau đó chính enzym này lại nối lại các đoạn bị đứt lại với nhau.



Hình 1.35. **Sao chép phân tử ADN mạch vòng theo dạng hình thêta (θ)**

Sao chép ADN ở sinh vật nhân chuẩn

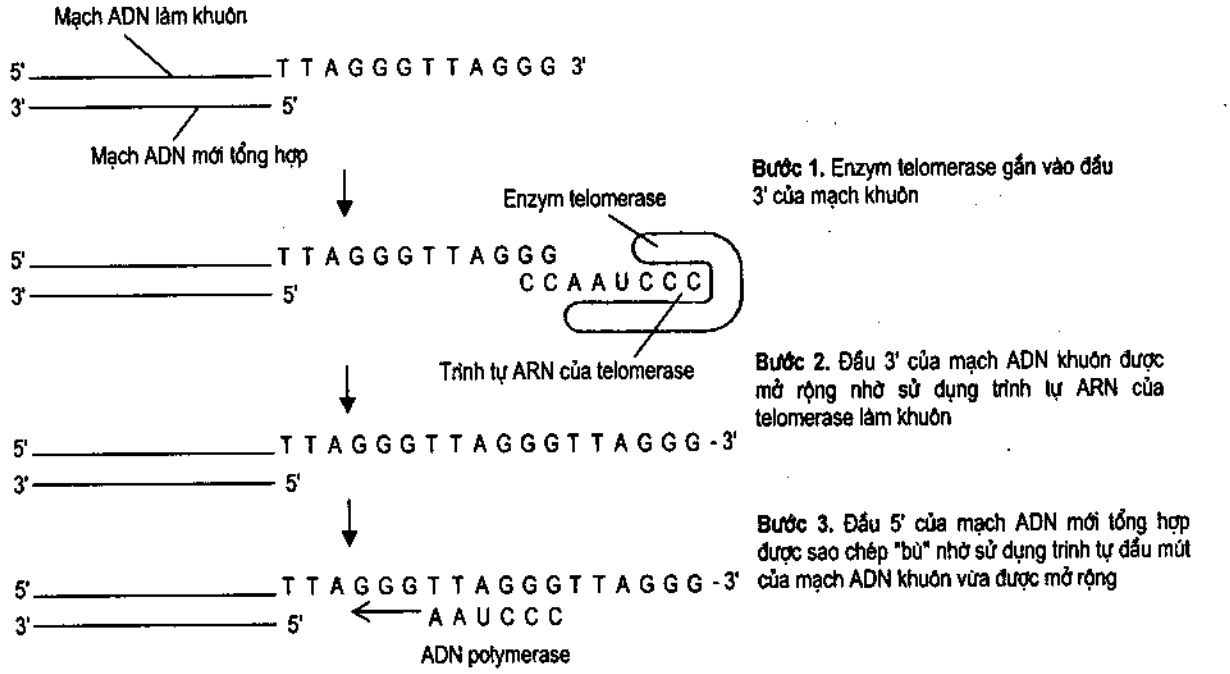
Trước khi tế bào phân chia, vật chất di truyền (ADN) của nó phải được sao chép. Sự phân chia tế bào ở sinh vật nhân chuẩn được điều hòa chặt chẽ và diễn ra qua nhiều pha trong chu trình tế bào (*Hình1.36*). Thời gian của một chu trình tế bào có thể biến động, song thường kéo dài trong nhiều giờ. Pha dài nhất là pha G_1, đây là pha tế bào chuẩn bị cho quá trình phân chia. Sau pha G_1 là pha S, là pha ADN được sao chép. Sau đó là một pha ngắn (pha G_2), rồi đến pha M là pha tế bào trải qua quá trình phân bào nguyên nhiễm (mitosis), bao gồm cả sự phân chia của nhiễm sắc thể và tế bào. Tiếp theo sau pha M, tế bào bước vào pha G_1 của chu trình tế bào kế tiếp. Cũng có khả năng là tế bào thoát khỏi chu trình tế bào và chuyển sang pha G_0, là lúc tế bào giữ nguyên trạng thái giống với giai đoạn nó đã đạt đến trong chu trình tế bào. Ở một số loại tế bào, như các nơron thần kinh, sau quá trình biệt hóa, tế bào hoàn toàn không phân chia nữa và tồn tại vĩnh viễn ở pha G_0.



Hình 1.36. **Chu trình tế bào ở sinh vật nhân chuẩn**

Do ADN của các sinh vật nhân chuẩn có kích thước lớn nên sự sao chép được khởi đầu cùng lúc từ nhiều điểm khởi đầu sao chép khác nhau nhằm đảm bảo cho quá trình sao chép ADN có thể hoàn thành trong một khoảng thời gian hợp lý. Chạc sao chép xuất phát từ mỗi điểm khởi đầu sao chép sẽ lan tỏa theo cả hai hướng tạo thành vòng sao chép, để cuối cùng các vòng sao chép gặp nhau và hòa hợp lại. Đoạn ADN được sao chép bắt đầu từ cùng một điểm khởi đầu sao chép được gọi là một đơn vị sao chép. Tế bào của động vật có vú điển hình có khoảng 50.000 - 100.000 đơn vị sao chép; mỗi đơn vị sao chép gồm khoảng 40 - 200 cặp base. Không phải tất cả các phân tử ADN đều được sao chép đồng thời. Cứ một nhóm gồm khoảng 50 đơn vị sao chép được sao chép đồng thời vào cùng một thời điểm trong pha S. Có xu hướng là các vùng ADN chứa những gen có hoạt tính phiên mã cao được sao chép trước, còn vùng chứa các trình tự không được phiên mã được sao chép sau.

ADN của các sinh vật nhân chuẩn được bao gói trong một phức hệ với protein gọi là các nucleosom (thể nhân). Trong quá trình sao chép ADN, phân tử ADN phải được tháo xoắn khỏi các nucleosom để cho quá trình sao chép có thể diễn ra. Điều này không chỉ tạo điều kiện cho chạc sao chép có thể tiến về phía trước mà còn góp phần giải thích cho sự hình thành các đoạn Okazaki ngắn (100 - 200 base) trên sợi theo sau. Sau khi chạc sao chép đi qua, cấu trúc của các nucleosom được khôi phục lại.



Hình 1.37. **Sao chép ADN ở phần đầu mút nhiễm sắc thể**

Việc sao chép ADN mạch thẳng của các sinh vật nhân chuẩn gặp phải một trở ngại không giống với sự sao chép ADN mạch vòng ở vi khuẩn. Đó là đầu 5' của các mạch ADN mới được tổng hợp không thể sao chép được vì không có đoạn mồi ARN phía trước để các enzym ADN polymerase có thể khởi đầu hoạt động tổng hợp ADN. Vì vậy, các nhiễm sắc thể sẽ bị ngắn dần đi sau mỗi lần sao chép và dẫn đến việc làm mất thông tin di truyền. Khó khăn này được khắc phục nhờ một cấu trúc đặc biệt ở đầu mút của nhiễm sắc thể. Phần đầu mút của nhiễm sắc thể thường chứa những đoạn trình tự không mã hóa, thành phần đơn giản và lặp lại liên

tiếp nhiều lần. Ở người, trình tự này là 5'-TTAGGG-3'. Enzym telomerase chứa phân tử ARN trong phân tử của nó có trình tự tương ứng là 3'-AAUCCC-5' có khả năng tự sử dụng trình tự này làm khuôn và tổng hợp đoạn ADN phần đầu mút và gắn vào mạch ADN làm khuôn giúp kéo dài mạch này. Sau đó, enzym telomerase tách ra và gắn với đầu mút của nhiễm sắc thể mới, nhờ vậy mà phần đầu nhiễm sắc thể được kéo dài ra. Sau đó, mạch làm khuôn sẽ được dùng để sao chép "bù" phần đầu 5' của mạch mới được tổng hợp (*Hình1.37*). Như vậy, hai quá trình nêu trên dẫn đến hiện tượng là ADN bị cắt ngắn sau mỗi lần sao chép bình thường, nhưng được kéo dài trở lại nhờ hoạt động của enzym telomerase.

I.K ĐIỀU HÒA BIỂU HIỆN CỦA GEN Ở SINH VẬT NHÂN SƠ

Điều hòa sự biểu hiện của gen

Để vi khuẩn hoạt động chức năng, không phải tất cả các gen trong hệ gen của nó đều cần được phiên mã như nhau. Thay vào đó, như một cơ chế bảo tồn năng lượng và nguyên liệu, vi khuẩn điều hòa sự biểu hiện của các gen sao cho chỉ những sản phẩm cần thiết cho hoạt động chức năng của tế bào mới được tạo ra. Chẳng hạn, sẽ vô ích nếu như vi khuẩn sản sinh ra các enzym cần cho quá trình tổng hợp các amino acid trong khi môi trường tế bào đã dư thừa nguồn nguyên liệu này. Nhờ quá trình điều hoà hoạt động của gen mà vi khuẩn có thể đáp ứng lại những thay đổi của môi trường, trong đó điển hình là việc đáp ứng lại việc môi trường có hoặc không có các chất dinh dưỡng thiết yếu.

Vi khuẩn điều hoà hoạt động gen nhằm kiểm soát lượng sản phẩm của gen trong tế bào. Lượng sản phẩm của một gen được coi là ổn định khi có sự cân bằng giữa tốc độ biểu hiện của gen đó (tổng hợp protein do gen đó mã hóa) và tốc độ phân hủy của sản phẩm của gen (phân tử protein đó). Trong thực tế, sự thay đổi tốc độ và mức độ biểu hiện gen là nguyên nhân trực tiếp làm thay đổi hàm lượng sản phẩm gen. Sự biểu hiện của gen có thể được điều hòa bởi các yếu tố và cơ chế khác nhau, trong đó có các cơ chế sau:

– Thay đổi tốc độ phiên mã.

– Thay đổi thời gian tồn tại của mARN.

– Thay đổi tốc độ dịch mã.

Trong thực tế, các cơ chế trên có thể cùng ảnh hưởng đến sự biểu hiện của gen nhưng sự điều hoà phiên mã là cơ chế điều hòa hoạt động của gen nổi bật hơn cả.

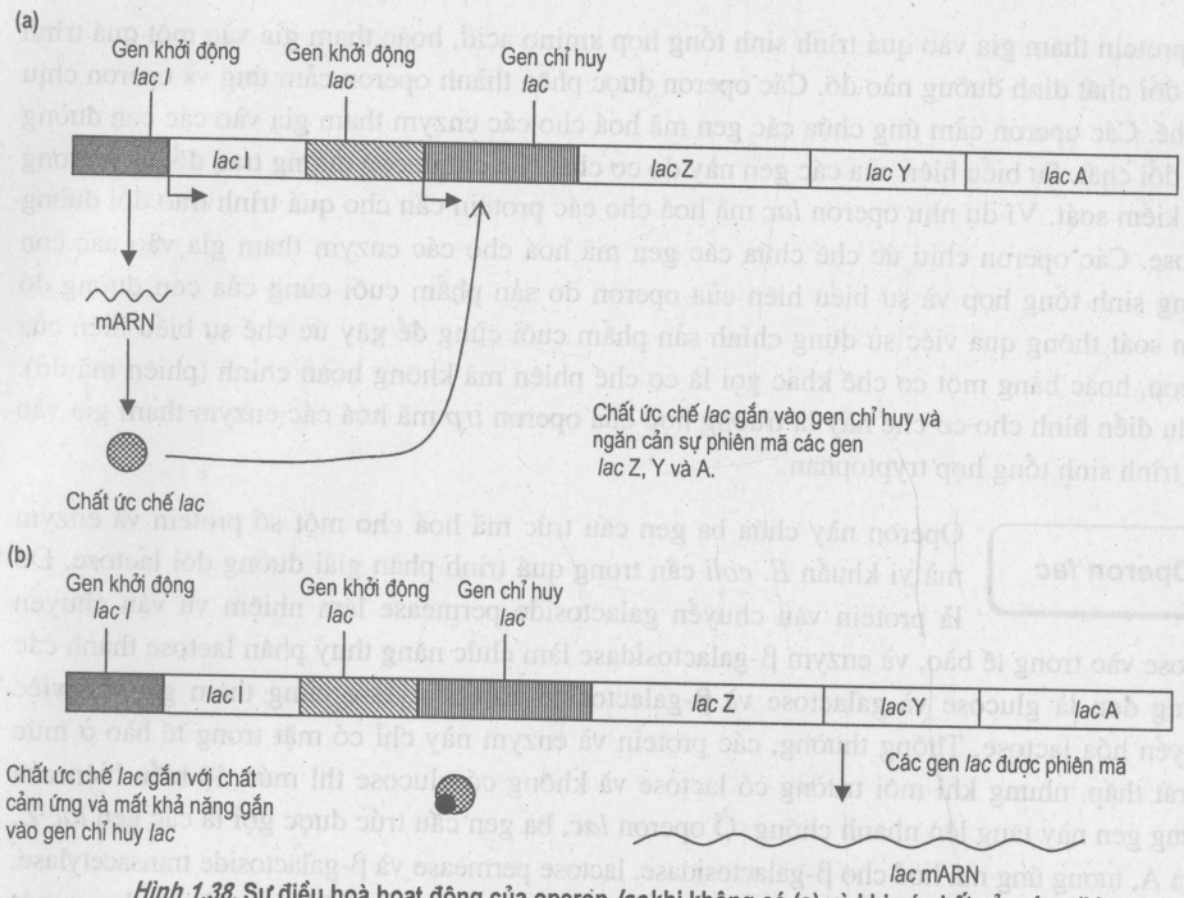
Tổ chức của các gen vi khuẩn

Đặc điểm quan trọng xác định cách thức điều hòa phiên mã ở vi khuẩn là sự tổ chức các gen thành các operon. Trong đó, các operon là các đơn vị phiên mã có một vài gen nằm liền kề, thường là các gen mã hoá cho các protein có chức năng liên quan với nhau được điều hòa biểu hiện đồng thời. Trong operon còn có thể có những gen khác, như gen mã hoá cho các protein điều hoà, để kiểm soát sự biểu hiện của chính operon. Ở *E. coli*, người ta đã phát hiện có nhiều operon khác nhau. Hầu hết các operon này chứa các gen mã hoá cho

các protein tham gia vào quá trình sinh tổng hợp amino acid, hoặc tham gia vào một quá trình trao đổi chất dinh dưỡng nào đó. Các operon được phân thành operon cảm ứng và operon chịu ức chế. Các operon cảm ứng chứa các gen mã hoá cho các enzym tham gia vào các con đường trao đổi chất. Sự biểu hiện của các gen này do cơ chất của chính con đường trao đổi chất tương ứng kiểm soát. Ví dụ như operon *lac* mã hoá cho các protein cần cho quá trình trao đổi đường lactose. Các operon chịu ức chế chứa các gen mã hoá cho các enzym tham gia vào các con đường sinh tổng hợp và sự biểu hiện của operon do sản phẩm cuối cùng của con đường đó kiểm soát thông qua việc sử dụng chính sản phẩm cuối cùng để gây ức chế sự biểu hiện của operon, hoặc bằng một cơ chế khác gọi là cơ chế phiên mã không hoàn chỉnh (phiên mã dở). Ví dụ điển hình cho cơ chế này là trường hợp của operon *trp* mã hoá các enzym tham gia vào quá trình sinh tổng hợp tryptophan.

Operon *lac*

Operon này chứa ba gen cấu trúc mã hoá cho một số protein và enzym mà vi khuẩn *E. coli* cần trong quá trình phân giải đường đôi lactose. Đó là protein vận chuyển galactoside permease làm nhiệm vụ vận chuyển lactose vào trong tế bào, và enzym β-galactosidase làm chức năng thuỷ phân lactose thành các đường đơn là glucose và galactose và β-galactoside transacetylase cũng tham gia vào việc chuyển hóa lactose. Thông thường, các protein và enzym này chỉ có mặt trong tế bào ở mức độ rất thấp, nhưng khi môi trường có lactose và không có glucose thì mức độ biểu hiện của những gen này tăng lên nhanh chóng. Ở operon *lac*, ba gen cấu trúc được gọi là các gen *lac* Z, Y và A, tương ứng mã hoá cho β-galactosidase, lactose permease và β-galactoside transacetylase. Các gen này nằm liền kề nhau và cùng được phiên mã thành một mARN bắt đầu từ chung một trình tự khởi động (promoter, còn gọi là gen khởi động). Ngoài ra, trong operon *lac* còn có một gen điều hoà khác (gọi là *lac* I) được biểu hiện riêng rẽ nằm ngược dòng operon và mã hoá cho protein được gọi là protein ức chế *lac* I làm nhiệm vụ điều hoà sự biểu hiện của các gen *lac* Z, Y và A. Khi trong môi trường không có lactose, protein ức chế *lac* gắn vào đoạn trình tự ADN gọi là trình tự chỉ huy (operator, còn gọi là gen chỉ huy) nằm giữa trình tự khởi động *lac* và điểm bắt đầu của gen *lac* Z. Khi gắn vào gen chỉ huy, chất ức chế *lac* khóa đường dẫn để gắn vào trình tự khởi động *lac* của enzym ARN polymerase và ngăn cản sự phiên mã các gen *lac*. Khi tế bào có lactose, một lượng nhỏ protein vận chuyển permerase vốn có trong tế bào cho phép lactose được chuyển vào trong tế bào chất. Lúc này allolactose - một đồng phân của lactose được sinh ra như một chất trung gian trong quá trình trao đổi chất lactose - được hình thành và đóng vai trò như một chất cảm ứng. Allolactose gắn vào protein ức chế đồng thời làm thay đổi cấu hình không gian của nó sao cho protein này không còn gắn được vào trình tự khởi động nữa. Nhờ vậy, đường dẫn cho enzym ARN polymerase không còn bị khoá và operon được phiên mã. Đồng thời lúc này, một lượng lớn protein galactoside permerase được sinh ra làm nhiệm vụ vận chuyển lactose vào tế bào và tham gia quá trình trao đổi chất. Như vậy, sự có mặt của lactose đã tác dụng cảm ứng sự biểu hiện của các enzym cần cho chính quá trình trao đổi chất của nó. Khi hết lactose trong môi trường, chất ức chế *lac* trở lại cấu hình ban đầu, gắn vào gen chỉ huy *lac* và ngăn cản sự phiên mã và operon bị đóng (*Hình 1.38*).

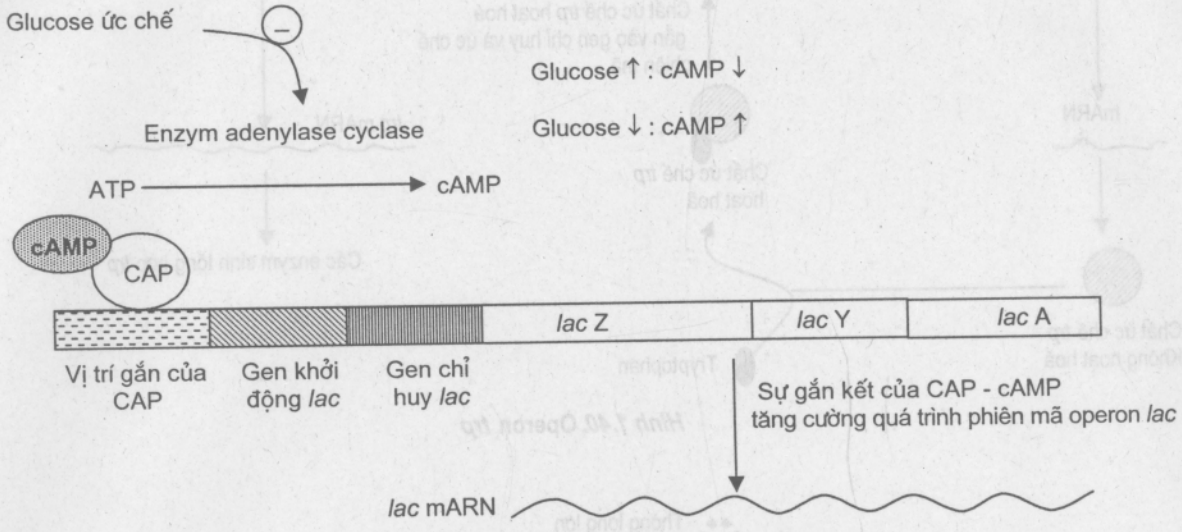


Hình 1.38. **Sự điều hoà hoạt động của operon *lac* khi không có (a) và khi có chất cảm ứng (b)**

Ức chế bởi chất dị hóa

Thuật ngữ này mô tả một cơ chế điều hòa bổ sung cho phép operon *lac* cảm nhận được sự có mặt của glucose vốn là nguồn năng lượng được ưu tiên hơn lactose. Nếu trong môi trường có cả lactose và glucose thì tế bào sử dụng glucose trước mà không cần tiêu phí năng lượng để phân cắt lactose thành các đường đơn. Sự có mặt của glucose trong tế bào sẽ ngăn cản sự biểu hiện của operon *lac* thông qua một cơ chế gọi là **sự ức chế bởi chất dị hoá** với sự tham gia của một protein điều hòa là **protein hoạt hoá chất dị hoá** CAP (*c*atabolite *a*ctivator *p*rotein, *Hình 1.39*). CAP khi được hoạt hóa có thể gắn một vị trí trên trình tự khởi động nằm ngược dòng operon *lac* và tăng cường sự liên kết của ARN polymerase vào trình tự khởi động dẫn đến tăng cường hoạt động phiên mã. Tuy nhiên, CAP chỉ gắn được vào trình tự khởi động khi có mặt (được hoạt hóa bởi) một dẫn xuất của ATP là adenosine monophosphat vòng (cyclic AMP), ký hiệu là cAMP. Lượng phân tử này có trong tế bào phụ thuộc vào lượng glucose có mặt trong môi trường. Enzym **adenylase cyclase** xúc tác việc hình thành cAMP và bị ức chế bởi glucose. Khi môi trường có glucose thì enzym adenylase cyclase bị ức chế và lượng cAMP được tạo ra thấp. Khi đó, CAP không gắn được vào trình tự khởi động và operon *lac* chỉ được phiên mã ở mức thấp. Ngược lại, khi môi trường có ít hoặc không có glucose, enzym adenylase cyclase không bị ức chế, lượng cAMP được tạo ra cao hơn và thúc đẩy CAP gắn vào trình tự khởi động, dẫn đến tốc độ phiên mã được tăng cường. Nhờ cơ chế như vậy, nên trong trường hợp tế bào có cả glucose và lactose thì operon *lac* cũng chỉ được phiên mã ở mức độ thấp. Khi glucose được sử dụng hết thì sự ức chế bằng chất dị hoá kết thúc và hoạt

động phiên mã operon *lac* được tăng cường giúp tế bào có thể sử dụng lactose làm nguồn cung cấp năng lượng chính.



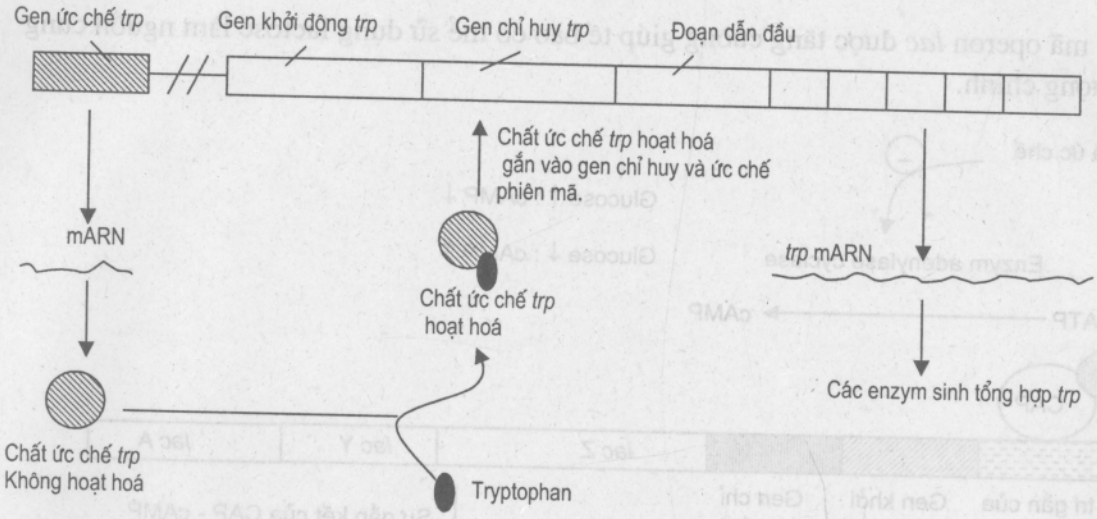
Hình 1.39. **Ức chế operon *lac* bằng chất dị hoá**

Operon *trp*

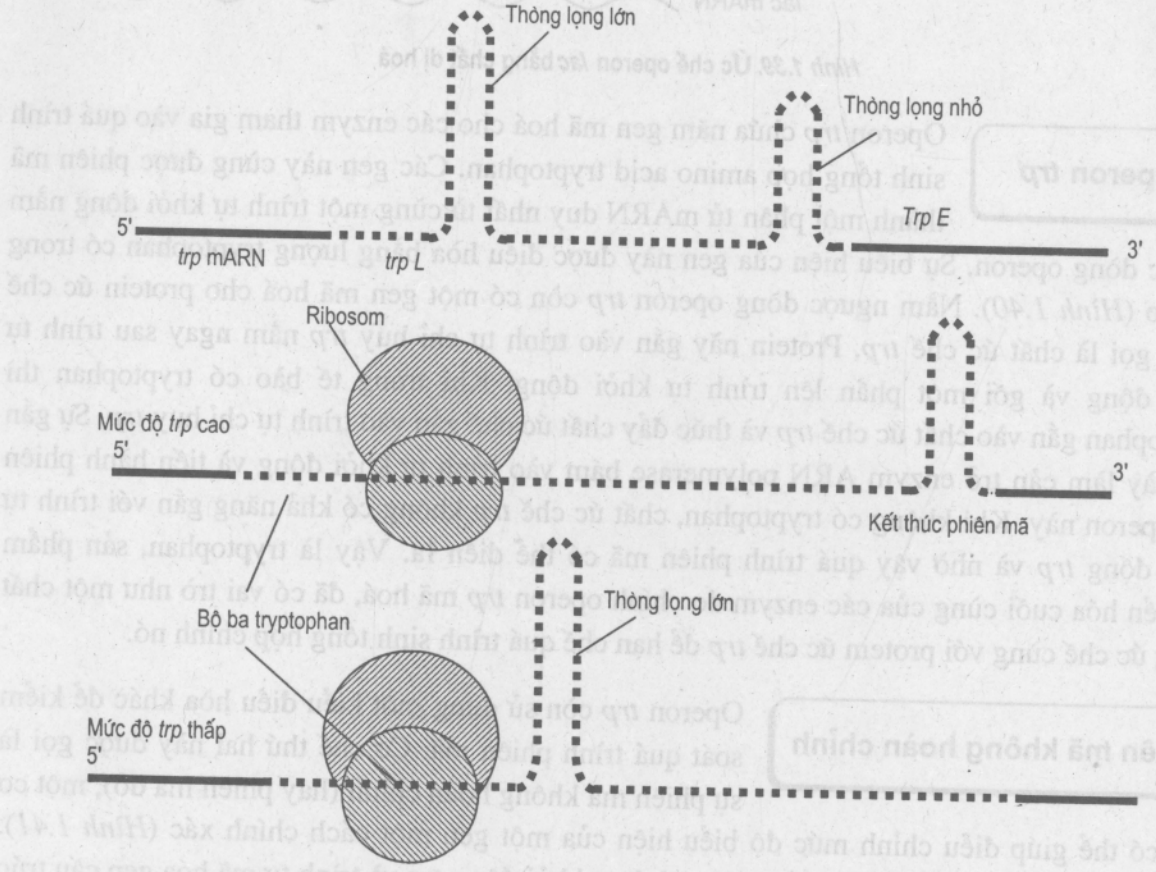
Operon *trp* chứa năm gen mã hoá cho các enzym tham gia vào quá trình sinh tổng hợp amino acid tryptophan. Các gen này cùng được phiên mã thành một phân tử mARN duy nhất từ cùng một trình tự khởi động nằm ngược dòng operon. Sự biểu hiện của gen này được điều hòa bằng lượng tryptophan có trong tế bào (*Hình 1.40*). Nằm ngược dòng operon *trp* còn có một gen mã hoá cho protein ức chế được gọi là chất ức chế *trp*. Protein này gắn vào trình tự chỉ huy *trp* nằm ngay sau trình tự khởi động và gối một phần lên trình tự khởi động. Khi trong tế bào có tryptophan thì tryptophan gắn vào chất ức chế *trp* và thúc đẩy chất ức chế gắn vào trình tự chỉ huy *trp*. Sự gắn kết này làm cản trở enzym ARN polymerase bám vào trình tự khởi động và tiến hành phiên mã operon này. Khi không có tryptophan, chất ức chế *trp* không có khả năng gắn với trình tự khởi động *trp* và nhờ vậy quá trình phiên mã có thể diễn ra. Vậy là tryptophan, sản phẩm chuyển hóa cuối cùng của các enzym do chính operon *trp* mã hoá, đã có vai trò như một chất đồng ức chế cùng với protein ức chế *trp* để hạn chế quá trình sinh tổng hợp chính nó.

Phiên mã không hoàn chỉnh

Operon *trp* còn sử dụng một kiểu điều hòa khác để kiểm soát quá trình phiên mã. Cơ chế thứ hai này được gọi là sự phiên mã không hoàn chỉnh (hay phiên mã dở), một cơ chế có thể giúp điều chỉnh mức độ biểu hiện của một gen một cách chính xác (*Hình 1.41*). Theo cơ chế này, đoạn mARN nằm giữa trình tự khởi động *trp* và trình tự mã hóa gen cấu trúc thứ nhất có khả năng tạo ra hai cấu trúc thòng lọng. Hai cấu trúc thòng lọng này không cùng được hình thành tại cùng một thời điểm mà chỉ có một trong hai loại này có thể xảy ra. Thòng lọng lớn có cấu trúc bền vững hơn nằm ngược dòng với thòng lọng nhỏ. Nếu thòng lọng lớn không làm ảnh hưởng đến sự phiên mã thì thòng lọng nhỏ đóng vai trò như một dấu hiệu kết thúc phiên mã.



Hình 1.40. **Operon *trp***



Hình 1.41. **Cơ chế phiên mã không hoàn chỉnh (phiên mã dở) ở operon *trp***

Khi thông lọng nhỏ được hình thành, sự phiên mã sẽ dừng lại trước gen *trp* cấu trúc đầu tiên. Sự phiên mã dở phụ thuộc vào sự kiện ở vi khuẩn là quá trình phiên mã và dịch mã cùng có thể xảy ra đồng thời. Lúc đó, các ribosom gắn vào mARN để tiến hành dịch mã ngay khi phân tử mARN đó còn đang được tổng hợp. Sự gắn kết của ribosom vào mARN chi phối việc loại thông lọng nào được hình thành và do đó đồng thời xác định việc quá trình phiên mã được

tiếp tục hay kết thúc. Phía trước thông lọng có một khung đọc mở ngắn chứa 14 bộ ba và một bộ ba kết thúc được dịch mã trước khi sự dịch mã các gen cấu trúc xảy ra. Trong số 14 bộ ba kể trên có hai bộ ba mã hóa cho tryptophan. Trong tế bào có dư thừa tryptophan thì ribosom sẽ dịch mã đoạn khung đọc mở này ngay sau khi ARN polymerase hoàn thành việc phiên mã đoạn phân tử mARN đó. Trong tình huống đó, sự có mặt của ribosom ngay sau phân tử ARN polymerase sẽ ngăn cản sự hình thành thông lọng lớn đồng thời dẫn đến sự hình thành thông lọng nhỏ và kết thúc quá trình phiên mã. Nếu tế bào thiếu tryptophan, ribosom sẽ ngừng lại tại vị trí các bộ ba mã hóa tryptophan. Lúc này, enzym ARN polymerase sẽ dịch chuyển lên phía trước để cho cấu trúc thông lọng lớn hình thành. Việc hình thành thông lọng nhỏ sẽ bị ngăn cản và quá trình phiên mã có thể tiếp diễn. Tốc độ ribosom dịch mã các bản sao có thể được điều khiển phụ thuộc vào nồng độ tryptophan có trong môi trường. Khi trong môi trường có lượng tryptophan trung bình, một số bản sao ngừng dịch mã, trong khi một số khác vẫn tiếp tục để điều chỉnh mức độ phiên mã của operon. Nói chung, chính chất ức chế *trp* xác định việc operon được đóng hay mở và cơ chế phiên mã dở điều chỉnh mức độ hoạt động của quá trình phiên mã (tinh chỉnh). Cả hai cơ chế này đều phụ thuộc vào lượng tryptophan có trong tế bào. Cơ chế phiên mã dở cho phép tế bào tổng hợp tryptophan tùy thuộc vào yêu cầu của nó. Cơ chế phiên mã dở không chỉ giới hạn đối với operon *trp* mà còn diễn ra ít nhất ở sáu operon mã hoá cho các enzym sinh tổng hợp các amino acid khác. Chẳng hạn như, một số operon như *trp* và *phe* được điều hòa bằng cả hai cơ chế: 1) bởi chất ức chế và 2) cơ chế phiên mã dở; trong khi một số operon khác như *his*, *leu* và *thr* chỉ được điều hòa theo cơ chế phiên mã dở.

Điều hòa bằng các yếu tố σ thay thế

Enzym ARN polymerase của vi khuẩn được cấu tạo từ năm chuỗi polypeptid (2 chuỗi α, 1 chuỗi β, 1 chuỗi β', và một chuỗi σ). Ở vi khuẩn, yếu tố sigma (chuỗi σ) chịu trách nhiệm khởi đầu quá trình phiên mã thông qua việc nhận biết đoạn trình tự khởi động của gen. Vi khuẩn, kể cả *E. coli*, có khả năng tổng hợp nên những yếu tố σ thay thế có thể nhận biết được nhiều trình tự khởi động khác nhau, nhờ đó ARN polymerase có thể phiên mã được nhiều gen khác nhau. Hiện tượng đó được sử dụng như một phương pháp để điều hoà sự biểu hiện của gen khi điều kiện môi trường đòi hỏi phải có sự thay đổi lớn về mức độ biểu hiện của gen. Yếu tố σ^{70} là yếu tố phổ biến nhất được tìm thấy ở *E. coli*. Các yếu tố σ thay thế hoạt động trong những tình huống sau:

– **Sốc nhiệt.** Khi tiếp xúc với nhiệt độ cao, *E. coli* phiên mã một nhóm gồm 17 protein giúp nó thích nghi với điều kiện môi trường thay đổi. Yếu tố σ^{12} thay thế được tạo ra để nhận biết các gen khởi động của các gen mẫn cảm với điều kiện sốc nhiệt.

– **Sinh bào tử ở *Bacillus subtilis*.** Vi khuẩn này sinh bào tử khi điều kiện môi trường bất lợi. Quá trình sinh bào tử đòi hỏi những thay đổi mạnh mẽ về sự biểu hiện của gen, bao gồm từ việc ngừng hầu hết các quá trình sinh tổng hợp protein và tái sản sinh các protein cần thiết để bắt đầu lại các quá trình đó khi bào tử nảy mầm. *B. subtilis* đã sử dụng yếu tố σ thay thế để thực hiện sự thay đổi đó.

– **Các yếu tố σ thay thế của bacteriophage.** Một số phage (phagơ) sử dụng ARN polymerase của tế bào vật chủ, nhưng dùng các yếu tố σ của mình để cải biến enzym nhằm có thể phiên mã các gen của phagơ một cách hoàn hảo (xem mục II.H). Một số loại phagơ khác lại sinh ra hàng loạt các yếu tố σ cho phép chúng kiểm soát sự biểu hiện của các gen riêng biệt của mình theo thời gian bằng cách bổ sung các yếu tố σ thay thế vào ARN polymerase của tế bào vật chủ làm nhiệm vụ phiên mã các gen biểu hiện sớm để các enzym này đồng thời được sử dụng để phiên mã các gen biểu hiện muộn hơn.

I.L ĐIỀU HÒA BIỂU HIỆN GEN Ở SINH VẬT NHÂN CHUẨN

Sự biểu hiện của gen nhân chuẩn

Hệ gen người ước tính chứa khoảng 30.000 - 40.000 gen, có các phương thức điều hoà hoạt động phức tạp, đến nay chưa được hiểu biết đầy đủ. Ở các tế bào nhân chuẩn, không phải tất cả các gen trong tế bào đều được biểu hiện đồng thời. Mỗi tế bào chỉ biểu hiện khoảng 15% số gen của nó; số còn lại không được biểu hiện. Ở sinh vật đa bào, những gen khác nhau được hoạt hóa biểu hiện ở những tế bào khác nhau. Sự biểu hiện khác nhau của các gen trong hệ gen quy định loại protein và enzym được tổng hợp trong tế bào, và qua đó xác định đặc điểm và chức năng đặc thù của tế bào trong cơ thể. Chẳng hạn như, trong các tế bào lympho làm nhiệm vụ sản sinh kháng thể chống lại sự nhiễm khuẩn thì các gen mã hóa các chuỗi polypeptid tạo nên kháng thể được biểu hiện ở mức độ cao. Hình thức biểu hiện của các gen có thể được thay đổi trong các giai đoạn khác nhau của vòng đời tế bào. Ví dụ, các tế bào máu được biệt hoá từ các tế bào tiền thân nguyên thuỷ. Những thay đổi đặc tính của tế bào khi nó biệt hóa là kết quả của sự thay đổi hình thức biểu hiện gen. Việc tìm hiểu hình thức điều hoà sự biểu hiện gen rất quan trọng trong việc tìm hiểu cơ chế gây nhiều chứng bệnh di truyền và các bệnh ung thư. Trong đó, sự biểu hiện bất thường của các gen có thể dẫn đến sự phân chia tế bào trong tình trạng không kiểm soát được, tạo thành khối u và dẫn đến ung thư.

Điều hòa phiên mã

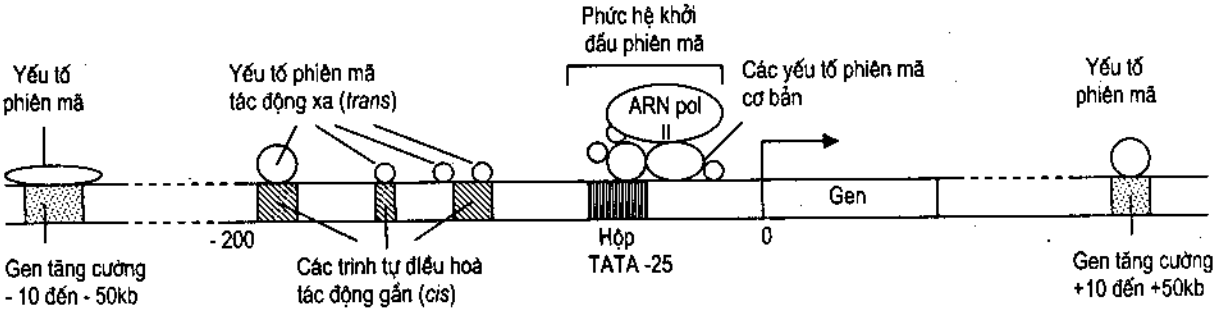
Các tế bào ở sinh vật nhân chuẩn có thể điều hoà sự biểu hiện của gen thông qua sự thay đổi tốc độ phiên mã. Sự điều hoà này có thể do sự tương tác giữa trình tự khởi động (trình tự khởi đầu phiên mã, promoter) với các protein liên kết với nó được gọi là các yếu tố phiên mã (*Hình 1.42*). Enzym ARN polymerase khởi đầu phiên mã tại vị trí trình tự khởi động, và hiệu quả của sự khởi đầu phiên mã này có quan hệ trực tiếp với sự liên kết của các yếu tố phiên mã vào một số vị trí đặc trưng của trình tự khởi động. Các đoạn trình tự tại các vị trí liên kết của protein điều hòa nằm trong đoạn trình tự khởi động cùng chiều với trình tự của gen cấu trúc nên được gọi là **cơ chế tác động gần (*cis*-acting).**

Ở các tế bào nhân chuẩn, tất cả các gen mã hóa protein đều được phiên mã nhờ enzym ARN polymerase II. Sự phiên mã được khởi đầu bằng phức hệ khởi đầu phiên mã (TIC) được hình thành do sự liên kết của enzym ARN polymerase II với một số phân tử protein đặc trưng được gọi là **các yếu tố phiên mã cơ bản**. Phức hệ này gắn vào trình tự khởi động của gen tại một đoạn trình tự đặc biệt gọi là **hộp TATA** (xem mục I.D). Hộp này có trình tự 5'-TATA(A/T)A(A/T)-3' và có mặt ở hầu hết (nhưng không phải toàn bộ) các gen, nằm phía trước (ngược dòng) vị trí khởi đầu phiên mã khoảng 25 cặp base (vị trí -25). Chức năng của hộp TATA là giúp định vị enzym ARN polymerase gắn vào đúng vị trí để khởi đầu phiên mã. Tuy vậy, có một số gen, đặc biệt là các gen chỉ biểu hiện ở các mô, tế bào chuyên biệt, không có hộp TATA. Thay vào đó, chúng có đoạn trình tự khởi động phiên mã thường nằm trùm lên điểm bắt đầu phiên mã. Một số gen khác, thường là các gen được biểu hiện ở mức thấp, không có cả cấu trúc hộp TATA lẫn cấu trúc đoạn trình tự khởi động nêu trên.

Hiệu quả khởi đầu phiên mã (tương ứng với lượng mARN sẽ được sinh ra) còn bị ảnh hưởng bởi các yếu tố phiên mã bổ sung liên kết với một số đoạn trình tự khác trong vùng trình

tự khởi động và tương tác với các protein của TIC làm thay đổi tính ổn định của phức hệ này. Các yếu tố phiên mã bổ sung có thể gây hiệu ứng tăng cường hoặc làm giảm tốc độ phiên mã. Các yếu tố phiên mã bổ sung có thể nhận biết và gắn vào một số vị trí của trình tự khởi động. Đoạn trình tự nhận biết có thể có độ dài khác nhau ở các gen khác nhau, và thường được mô tả là đoạn **trình tự liên ứng** có khả năng liên kết được với nhiều loại yếu tố phiên mã. Các yếu tố phiên mã được tổng hợp trong tế bào chất nhưng hoạt động trong nhân tế bào. Vì thế mà cơ chế tác động của chúng được coi là các **cơ chế tác động xa (*trans* - acting).**

Quá trình phiên mã cũng có thể bị ảnh hưởng bởi các trình tự tăng cường (enhancer) thường nằm cách điểm khởi đầu phiên mã của gen đó hàng nghìn cặp base. Các trình tự tăng cường thường có kích thước khoảng 100 - 200 cặp base, mang các đoạn trình tự liên kết được với các yếu tố phiên mã nên có thể thúc đẩy sự phiên mã của các gen có quan hệ với nó. Trình tự tăng cường có thể nằm ngược dòng hoặc xuôi dòng so với gen chịu ảnh hưởng của nó. Các trình tự tăng cường hoạt động độc lập, và chúng có hiệu quả như nhau dù có thể nằm ở các vị trí khác nhau. Sự tương tác giữa trình tự tăng cường và trình tự khởi động của một gen có thể được thực hiện bằng cách đoạn ADN nằm giữa chúng tạo cấu trúc thòng lọng để chúng tiến lại gần nhau. Một số trình tự điều hòa có chứa những đoạn liên kết được với các yếu tố cản trở phiên mã. Các trình tự đó còn được gọi là các trình tự gây câm (gen gây câm, silencer). Các trình tự này có thể làm hạn chế sự biểu hiện của gen trong một số loại tế bào chuyên biệt. Trong tế bào, còn có thể tồn tại những đoạn trình tự ở vị trí xa gen hơn, gọi là các yếu tố kiểm soát locut. Các yếu tố này có ảnh hưởng đến sự biểu hiện của toàn bộ một họ gen nào đó bằng cách kiểm soát sự tiếp cận của các yếu tố phiên mã đến đoạn trình tự ADN của họ gen đó. Chẳng hạn như, người ta đã phát hiện thấy các vùng kiểm soát locut của họ gen mã hóa cho các phân tử globin.



Hình 1.42. **Điều hoà sự biểu hiện của gen ở sinh vật nhân chuẩn**

Các yếu tố phiên mã

Các yếu tố phiên mã là một họ protein lớn có vai trò điều hòa sự biểu hiện của các gen mã hóa cho protein. Chúng khác biệt hẳn so với các yếu tố phiên mã cơ bản có chức năng tương tác với enzym ARN polymerase II để hình thành nên phức hệ khởi đầu phiên mã TIC. Các yếu tố phiên mã có nhiều hình thức hoạt động khác nhau, chẳng hạn một số yếu tố chỉ có ở những loại tế bào chuyên biệt trong khi một số khác có mặt ở tất cả các loại tế bào. Khi gắn vào trình tự khởi động, các yếu tố phiên mã có thể điều hoà sự phiên mã theo chiều hướng dương tính (thúc đẩy) hoặc âm tính (ức chế). Vai trò tổng thể đối với sự phiên mã phụ thuộc vào sự bù trừ các tác động thúc đẩy hay ức chế của các yếu tố phiên mã gắn vào. Khả năng tự tạo các liên kết kép (dimer) giữa các yếu tố phiên mã cùng loại hay với các yếu tố phiên mã khác nhau đã tạo cho quá trình điều hòa hoạt động của gen có thêm nhiều công cụ điều hòa khác nhau.

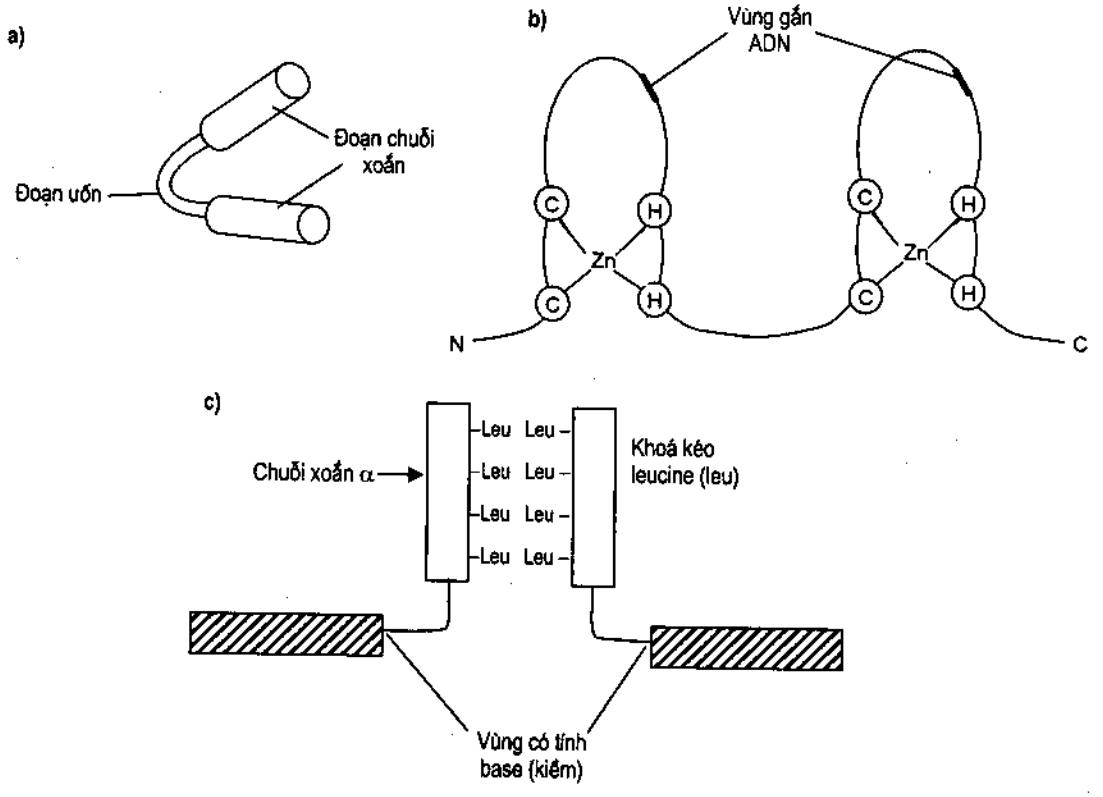
Các yếu tố phiên mã có cấu trúc modul gồm các miền protein đặc biệt với những chức năng đặc trưng. Trong đó, có ba kiểu miền cấu trúc cơ bản là:

– Miền liên kết ADN.

– Miền tạo liên kết kép (dimer hóa).

– Miền hoạt hoá xa.

Miền gắn kết ADN và miền tạo liên kết kép có chứa những cấu trúc protein đặc thù gọi là các motif (*Hình 1.43*).



Hình 1.43. **Các motif yếu tố phiên mã. (a) kiểu xoắn-uốn-xoắn, (b) kiểu ngón tay kẽm, và (c) Khoá kéo leucine và vùng gắn kết ADN có tính kiềm của yếu tố phiên mã Zip.**

Miền liên kết ADN:

Dựa vào đặc điểm các motif, người ta phân loại ba kiểu miền gắn kết ADN:

– Kiểu xoắn - uốn - xoắn. Motif này gồm hai vòng xoắn α tách biệt nhau bởi một mạch uốn kiểu β. Một trong số hai vòng xoắn đó được gọi là vòng xoắn nhận biết có vai trò gắn vào phân tử ADN thông qua tiếp xúc với khe lớn của chuỗi xoắn kép. Một ví dụ về các protein điều hòa phiên mã có cấu trúc motif này là họ các yếu tố phiên mã tham gia điều hòa hoạt động gen trong quá trình phát triển của phôi. Gen mã hóa chúng có tính bảo thủ cao.

– Ngón tay kẽm (Zn). Kiểu motif gắn kết ADN này tồn tại ở hai dạng là C_2H_2 và C_4. Dạng C_2H_2 hình thành một cấu trúc thòng lọng gồm 12 amino acid định vị vào một đế gồm hai cystein và hai histidin cùng được điều hòa xung quanh một ion kẽm (Zn). Cấu trúc đó tạo cho motif này có hình dạng giống ngón tay (vì vậy, gọi là ngón tay kẽm). Dạng motif này tạo nên một cấu trúc vững chắc gồm hai sợi xoắn β và một sợi xoắn α có chứa các amino acid

kiểm tương tác với ADN thông qua khe lớn của chuỗi xoắn kép. Motif ngón tay kẽm được lặp lại nhiều lần trong vùng liên kết ADN của yếu tố phiên mã và thường cần nhiều hơn 3 cấu trúc motif này để yếu tố phiên mã có thể gắn kết vào ADN. Có thể thấy cấu trúc motif này ở yếu tố phiên mã *Sp1*, một yếu tố phiên mã phổ biến. Motif C_4 có cấu trúc tương tự C_2H_2, nhưng ở đây bốn amino acid đều là cystein liên kết xung quanh một nguyên tử kẽm. Loại motif này tìm thấy ở các yếu tố phiên mã các gen mã hoá cho các thụ thể của các hormon steroid.

– **Miền có tính kiềm (miền base)**. Kiểu miền liên kết này thường tìm thấy có liên quan với một trong hai miền liên kết kép (dimer hóa) được gọi là cấu trúc khóa kéo leucin *bZIP* (*b*asic leucine *zip*per) do các amino acid leucin có kiềm tính tạo nên, và cấu trúc xoắn - uốn - xoắn HLH (helix - loop - helix) do các protein tạo cấu trúc HLH có kiềm tính tạo nên. Việc liên kết của yếu tố phiên mã vào phân tử ADN cần có hai miền base được đặt gần nhau nhờ sự hình thành liên kết kép (dimer hoá).

Miền dimer hóa:

Có hai kiểu motif được tìm thấy ở miền dimer hoá:

– **Motif khoá kéo leucin** thường có ở đầu tận cùng carboxyl của miền gắn kết ADN. Motif này chứa một chuỗi xoắn α, trong đó cứ bảy amino acid lại có amino acid leucin, nhờ vậy các tiểu phần amino acid leucin sẽ nằm quay về cùng một phía và có mặt cứ sau hai vùng cuộn của chuỗi xoắn, tạo nên một bề mặt kỵ nước của phân tử protein. Sự dimer hoá được tạo ra nhờ sự tương tác giữa những bề mặt kỵ nước của hai cấu trúc khoá kéo leucin và dẫn đến việc các miền liên kết ADN có tính kiềm của hai phân tử protein (yếu tố phiên mã) được đặt cạnh nhau. Do hai vùng liên kết ADN của hai phân tử ngược chiều nhau, nên chúng gắn với đoạn trình tự ADN theo kiểu đối xứng ngược chiều.

– **Vùng dimer hoá HLH** chứa hai chuỗi xoắn α được ngăn cách bởi một đoạn uốn không có cấu trúc chuỗi xoắn. Sự dimer hoá xảy ra nhờ sự tương tác giữa các amino acid kỵ nước cùng nằm ở đầu carboxyl của chuỗi xoắn. Sự dimer hoá có thể xảy ra không chỉ giữa hai phân tử của cùng một yếu tố phiên mã (được gọi là homodimer) mà có thể giữa các yếu tố phiên mã khác nhau (được gọi là heterodimer). Sự hình thành các heterodimer làm cho các yếu tố phiên mã có thể có những chức năng điều hòa phiên mã mới, đồng thời làm tăng khả năng điều hoà phiên mã các gen đích. Ví dụ về cơ chế điều hòa này là họ các yếu tố phiên mã *MyoD* có khả năng hình thành cả dạng homodimer và heterodimer tham gia điều hoà sự biểu hiện của gen ở các tế bào cơ.

Miền hoạt hoá:

Không giống như miền liên kết ADN và miền dimer hóa, đến nay không thấy có cấu trúc motif đặc trưng cho miền hoạt hóa.Việc phân tích trình tự amino acid chỉ cho thấy rằng miền hoạt hoá thường giàu một số loại amino acid nhất định. Trong đó, đặc biệt người ta đã xác định được nhiều miền hoạt hóa giàu thành phần các amino acid có tính acid (như yếu tố phiên mã *Gal4* ở nấm men), hoặc giàu glutamin (như yếu tố phiên mã *Sp1*), hoặc giàu prolin (như các yếu tố phiên mã *c-Jun Ap2* và *Oct-2*). Các yếu tố phiên mã điều hòa hoạt động của các gen đích bằng cách liên kết với phức hệ TIC và làm thay đổi tính ổn định của nó. Miền hoạt hoá cũng có thể tương tác với các protein khác nhau trong TIC vào các thời điểm khác nhau của giai đoạn khởi đầu và kéo dài phiên mã.

Một số yếu tố phiên mã có thể có vai trò ức chế quá trình phiên mã. Điều đó có thể thực hiện được bằng cách thay đổi hình thức điều hòa phiên mã khác nhau. Một số yếu tố có thể ức chế sự phiên mã bằng cách trực tiếp tương tác với phức hệ khởi đầu phiên mã. Các yếu tố khác có thể tác động một cách gián tiếp theo một trong những hình thức như sau: (i) ngăn cản sự liên kết của yếu tố thúc đẩy phiên mã vào vị trí liên kết trên phân tử ADN; (ii) hình thành dimer làm bất hoạt vùng liên kết ADN; hoặc (iii) gắn một protein ức chế vào vùng hoạt hoá của một yếu tố phiên mã.

Điều hòa hoạt động gen nhờ hormon và các hợp chất cytokin

Các hormon là các hợp chất do các tế bào chuyên biệt sinh ra có vai trò và chức năng tác động đến hoạt động của các tế bào khác thông qua sự hoạt hoá phiên mã các gen đặc biệt ở các tế bào đích. Các hormon có thể là những phân tử nhỏ, thường là steroid, như estrogen và glucocorticoid, hoặc là các chuỗi polypeptid (protein) như insulin. Các cytokin là nhóm các phân tử protein có tác động tương tự như hormon, với tế bào đích thông thường là các tế bào máu. Các hormon và cytokin điều hòa sự biểu hiện của gen ở các tế bào đích theo nhiều cách khác nhau. Chẳng hạn như, các hormon steroid có khả năng tan trong lipid và vì thế chúng có thể thâm nhập qua màng tế bào vào tế bào chất. Ở đó, chúng tương tác với các yếu tố phiên mã được gọi là các **thụ thể hormon steroid**. Việc tương tác này làm cho các thụ thể hormon steroid được giải phóng ra khỏi phức hệ protein ức chế. Sau đó, nó được dimer hóa và được vận chuyển vào nhân tế bào, rồi gắn vào đoạn trình tự khởi động và hoạt hóa sự phiên mã của các gen đích. Khác với các hormon steroid, các hormon có bản chất protein và cytokin có cơ chế tác động khác. Chúng thường gắn kết với protein thụ thể trên bề mặt của tế bào đích. Sự hoạt hoá của gen đích được thông qua một quá trình được gọi là quá trình **truyền tín hiệu**. Trong quá trình này, một loạt các phân tử protein lần lượt được hoạt hoá nhờ quá trình phosphoryl hóa. Kết quả cuối cùng là quá trình truyền tín hiệu kích thích sự phiên mã của các gen đích bằng cách tăng cường việc gắn kết của các yếu tố phiên mã với đoạn trình tự khởi động của gen.

Phần II
NHIỄM SẮC THỂ VÀ HỆ GEN

Phần II gồm các mục:

II.A Nhiễm sắc thể

Nhiễm sắc thể sinh vật nhân sơ và nhân chuẩn
Hình thái nhiễm sắc thể
Cấu trúc nhiễm sắc thể điển hình
Cấu trúc phân tử của nhiễm sắc thể
Chất nhiễm sắc hoạt động và không hoạt động chức năng
Sự biến đổi số lượng nhiễm sắc thể

II.B Sự phân chia tế bào (nguyên phân)

Chu trình tế bào ở sinh vật nhân chuẩn
Nguyên phân và sự phân chia tế bào chất
Sự điều hòa chu trình tế bào

II.C Hệ gen sinh vật nhân sơ

Nhiễm sắc thể sinh vật nhân sơ
Tổ chức hệ gen sinh vật nhân sơ
Plasmid
Gen nhảy (transposon) ở vi khuẩn
Vi khuẩn cổ

II.D Hệ gen người

Hệ gen người
Gen ở người
Họ gen
Gen giả
ADN liên gen
Các đoạn trình tự lặp lại phân tán
Các đoạn trình tự lặp lại co cụm
Các đoạn trình tự lặp lại có số lượng biến động (VNTR)

II.E Đột biến ADN

Đột biến
Đột biến điểm
Đột biến thô
Đột biến và bệnh di truyền
Đột biến ở mức độ cơ thể

II.F Các tác nhân gây đột biến và sửa chữa ADN

Các tác nhân gây đột biến
Các hóa chất gây đột biến
Các tác nhân vật lý gây đột biến
Sửa chữa ADN

II.G Tái tổ hợp

Tổng quan
Tái tổ hợp chung
Cấu trúc Holliday
Mô hình đứt gãy sợi kép
Sự biến trạng
Các enzym tái tổ hợp
Tái tổ hợp tại điểm đặc hiệu

II.H Phage

Phage
Chu trình sống của phage
Con đường gây tan của phage T4
Chu trình tiềm tan của phage lambda (λ)
Phage M13
Sự biểu hiện gen trong chu trình gây tan

II.I Virus lây nhiễm sinh vật nhân chuẩn

Virus
Hệ gen virus
Cấu tạo của virus
Các chiến lược sao chép
Virus ADN
Virus ARN
Retrovirus

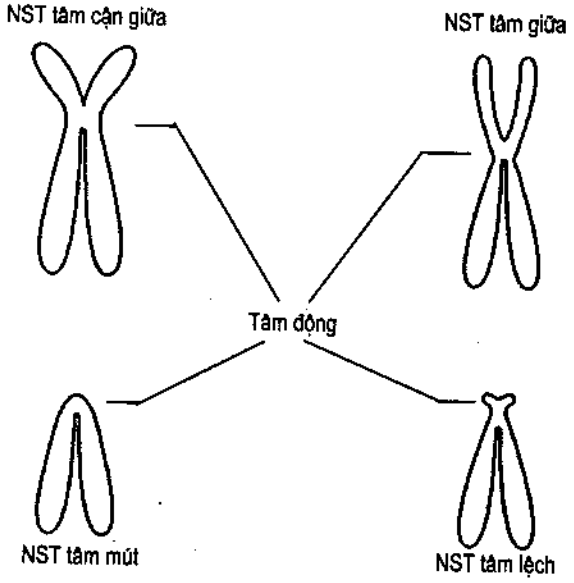
II.A NHIỄM SẮC THỂ

Nhiễm sắc thể sinh vật nhân sơ và nhân chuẩn

Tất cả các dạng sống có cấu tạo tế bào đều có cấu trúc vật chất di truyền là các nhiễm sắc thể (NST) mang các gen được mã hoá trên phân tử ADN. Tuy vậy, cấu trúc NST của các sinh vật nhân sơ và sinh vật nhân chuẩn rất khác nhau. Ở sinh vật nhân sơ, NST gồm một chuỗi ADN xoắn kép duy nhất, thường là mạch vòng liên kết với một số ít protein. Sự sao chép ADN diễn ra từ một điểm khởi đầu sao chép duy nhất. Trong khi đó, ở hầu hết các sinh vật nhân chuẩn, tế bào chứa nhiều NST có cấu trúc mạch thẳng và nằm trong một bào quan có màng bao bọc gọi là nhân tế bào. Các phân tử ADN liên kết chặt chẽ với một lượng lớn các phân tử protein đặc biệt. Các protein này có thể có vai trò trong cấu trúc hoặc chức năng của NST. Lượng ADN có trong tế bào ở sinh vật nhân chuẩn lớn hơn nhiều ở sinh vật nhân sơ. Vì thế, để đáp ứng yêu cầu về tốc độ sao chép, các NST ở sinh vật nhân chuẩn thường có nhiều điểm khởi đầu sao chép.

Hình thái nhiễm sắc thể

Các NST sinh vật nhân chuẩn thường chỉ nhìn thấy được khi tế bào đang phân chia, sau khi NST đã sao chép tạo thành một cấu trúc kép giống hệt nhau được gọi là **các nhiễm sắc tử** (các NST con). Các NST được phân loại dựa trên hình thái của nó. Hình thái đó được xác định bởi vị trí của tâm động (còn gọi là eo sơ cấp). *Hình 2.1* thể hiện bốn kiểu hình thái NST khác nhau. Ở các NST **tâm giữa**, tâm động nằm gần với điểm giữa của NST. Nó phân chia NST thành hai nửa bằng nhau (gọi là hai cánh, hoặc hai vai). Khi tâm động cách xa điểm giữa đủ để phân biệt được cánh dài (cánh *q*) và cánh ngắn (cánh *p*) thì NST được gọi là NST **tâm cận giữa**. Hai loại NST khác có tâm động gần về một đầu của NST. Ở NST **tâm mút**, tâm động ở tận cùng đầu NST, lúc này NST chỉ có một cánh. Nếu tâm động gần một đầu của NST tới mức cánh ngắn chỉ rất nhỏ thì NST được gọi là NST **tâm lệch**. Sự khác biệt giữa hai loại NST này dễ dàng nhận thấy trong trường hợp các NST ở người.



Hình 2.1. **Các dạng hình thái NST được phân loại dựa trên vị trí tâm động**

Một bộ NST lưỡng bội hoàn chỉnh của một loài được gọi là **kiểu nhân**. Các NST thường được đánh số thứ tự theo trật tự giảm dần về kích thước, riêng NST giới tính được ký hiệu là X hoặc Y. Trong thực nghiệm, có thể chụp hình được các NST, xác định được các cặp NST tương đồng và sắp xếp chúng theo trật tự, tạo thành **sơ đồ kiểu nhân**.

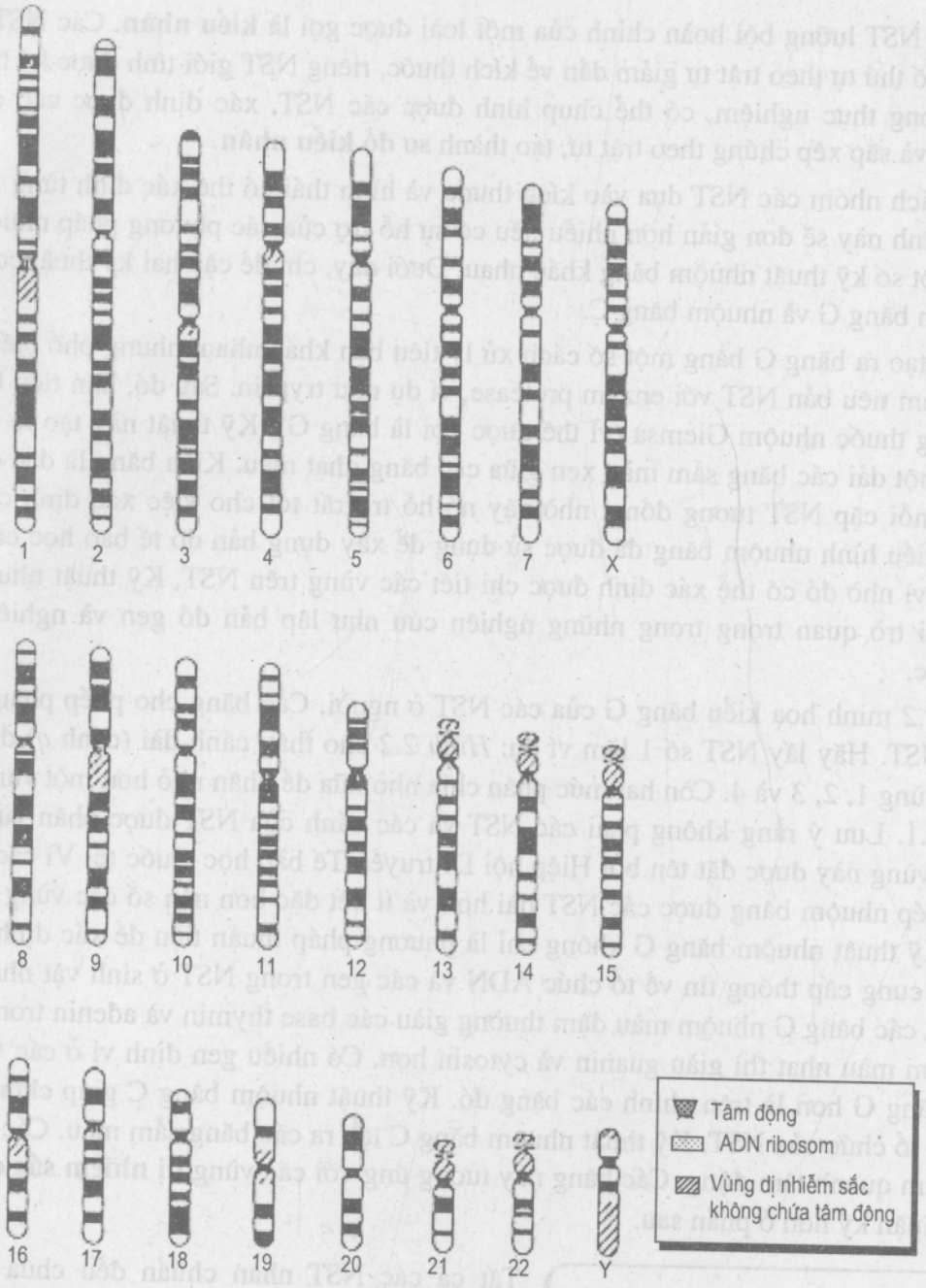
Bằng cách nhóm các NST dựa vào kích thước và hình thái có thể xác định từng NST của loài. Quá trình này sẽ đơn giản hơn nhiều nếu có sự hỗ trợ của các phương pháp nhuộm băng NST. Có một số kỹ thuật nhuộm băng khác nhau. Dưới đây, chỉ đề cập hai kỹ thuật cơ bản: kỹ thuật nhuộm băng G và nhuộm băng C.

Có thể tạo ra băng G bằng một số cách xử lý tiêu bản khác nhau, nhưng phổ biến nhất là xử lý nhẹ lam tiêu bản NST với enzym protease, ví dụ như trypsin. Sau đó, lam tiêu bản được nhuộm bằng thuốc nhuộm Giemsa (vì thế được gọi là băng G). Kỹ thuật này tạo ra hình ảnh NST gồm một dải các băng sẫm màu xen giữa các băng nhạt màu. Kiểu băng là đặc điểm đặc trưng cho mỗi cặp NST tương đồng, nhờ vậy nó hỗ trợ rất tốt cho việc xác định chính xác NST. Các kiểu hình nhuộm băng đã được sử dụng để xây dựng bản đồ tế bào học các NST ở nhiều loài, vì nhờ đó có thể xác định được chi tiết các vùng trên NST. Kỹ thuật nhuộm băng cũng có vai trò quan trọng trong những nghiên cứu như lập bản đồ gen và nghiên cứu di truyền y học.

Hình 2.2 minh hoạ kiểu băng G của các NST ở người. Các băng cho phép phân biệt các vùng của NST. Hãy lấy NST số 1 làm ví dụ: *Hình 2.2* cho thấy cánh dài (cánh *q*) được phân thành các vùng 1, 2, 3 và 4. Còn hai mức phân chia nhỏ nữa để phân nhỏ hơn một vùng cụ thể, ví dụ 1q42.1. Lưu ý rằng không phải các NST và các cánh của NST được phân vùng giống nhau. Các vùng này được đặt tên bởi Hiệp hội Di truyền Tế bào học Quốc tế. Vì các kỹ thuật này cho phép nhuộm băng được các NST dài hơn và ít kết đặc hơn nên số các vùng trên NST tăng lên. Kỹ thuật nhuộm băng G không chỉ là phương pháp thuận tiện để xác định NST mà còn có thể cung cấp thông tin về tổ chức ADN và các gen trong NST ở sinh vật nhân chuẩn. Nói chung, các băng G nhuộm màu đậm thường giàu các base thymin và adenin trong khi các vùng nhuộm màu nhạt thì giàu guanin và cytosin hơn. Có nhiều gen định vị ở các vùng nằm giữa các băng G hơn là trên chính các băng đó. Kỹ thuật nhuộm băng C giúp chúng ta hiểu sâu hơn về tổ chức của NST. Kỹ thuật nhuộm băng C tạo ra các băng sẫm màu. Các băng này chủ yếu nằm quanh tâm động. Các băng này tương ứng với các vùng **dị nhiễm sắc cơ yếu**, sẽ được thảo luận kỹ hơn ở phần sau.

Cấu trúc nhiễm sắc thể điển hình

Tất cả các NST nhân chuẩn đều chứa hai vùng khác biệt có tầm quan trọng đặc biệt về cấu trúc. Đó là tâm động và đầu mút. Ngoài ra, một số NST còn chứa vùng tổ chức hạch nhân. Tâm động là vị trí để các sợi thoi vô sắc gắn vào trong quá trình phân chia tế bào. Bất kỳ đoạn NST nào không liên kết với tâm động cũng sẽ không được phân ly về các tế bào con. Tâm động của nấm men gồm khoảng 200 cặp base là cấu trúc tâm động đã được nghiên cứu kỹ nhất. Hầu hết các tâm động ở các sinh vật nhân chuẩn khác đều có kích thước lớn hơn tâm động ở nấm men. Thường thì tâm động gồm các ADN vệ tinh có trình tự lặp lại cao. Ở người, có thể phân biệt các NST khác nhau nhờ vào các cấu trúc ADN vệ tinh đặc trưng ở tâm động. Sự liên kết của tâm động với các sợi thoi vô sắc là nhờ các protein gắn với tâm động tạo thành một cấu trúc gồm nhiều lớp gọi là **hạch phân chia** (thể động).



Hình 2.2. **Kiểu nhân ở người.** *Hình ảnh này minh họa hình thái nhiễm sắc thể (NST) trong 23 cặp NST tương đồng được nhuộm băng G (xem phần diễn giải). Các NST được xếp theo thứ tự kích thước giảm dần, bắt đầu từ NST số 1. NST X có kích thước nằm giữa NST số 7 và số 8. NST Y được xếp cuối cùng. Tâm động chia mỗi NST thành 2 cánh (vai). Dựa vào vị trí tâm động, người ta chia những NST này thành các nhóm hình thái khác nhau. Chỉ có NST số 2 và số 9 là NST tâm giữa. Các NST số 13, 14, 15, 21, 22 và Y là các NST tâm lệch. Còn lại là các NST tâm cận giữa với tỷ lệ giữa hai cánh biến động. Các NST 13, 14, 15, 21 và 22 chứa các trình tự vệ tinh mã hóa cho các ribosom (rADN).*

Các đầu mút NST không chỉ đơn giản là phần cuối của các NST hay phân tử ADN mà là một cấu trúc được biệt hóa. Chúng chứa nhiều đoạn trình tự ADN đơn giản, ngắn và lặp lại nhiều lần. Ở người, đoạn trình tự lặp lại đó là TTAGGG, nhưng có sự biến động nhỏ giữa các loài sinh vật nhân chuẩn khác nhau. Những đoạn trình tự giống nhau ở phần đầu mút được tìm thấy ở thực vật và các loài động vật nguyên sinh. Các protein đặc biệt liên kết ở vùng đầu mút

và tạo nên cấu trúc nucleoprotein. Người ta cho rằng cấu trúc đó có tác dụng cản trở sự tái tổ hợp giữa phần đầu của các NST khác nhau. Số lượng các đoạn lặp lại ở đầu mút thường nhiều ở các tế bào mầm nhưng giảm dần theo số lần phân bào ở các tế bào soma. Vì vậy, đây chính là chỉ thị phân tử xác định quá trình già hóa. Độ dài của đầu mút được duy trì nhờ hoạt động của một enzym đặc biệt gọi là **telomerase** - một loại protein có chứa trình tự ARN bổ trợ với đoạn ADN lặp lại tại đầu mút. Khi enzym này hoạt động, trình tự ARN được sử dụng như mạch khuôn để kéo dài phần đầu mút bị ngắn đi sau mỗi lần phân bào. Ở các tế bào soma thường vắng mặt enzym telomerase, nhưng enzym này xuất hiện ở nhiều dạng tế bào khối u. Ở những dòng tế bào khối u này (còn gọi là các dòng tế bào bất tử), độ dài phần đầu mút được duy trì ổn định.

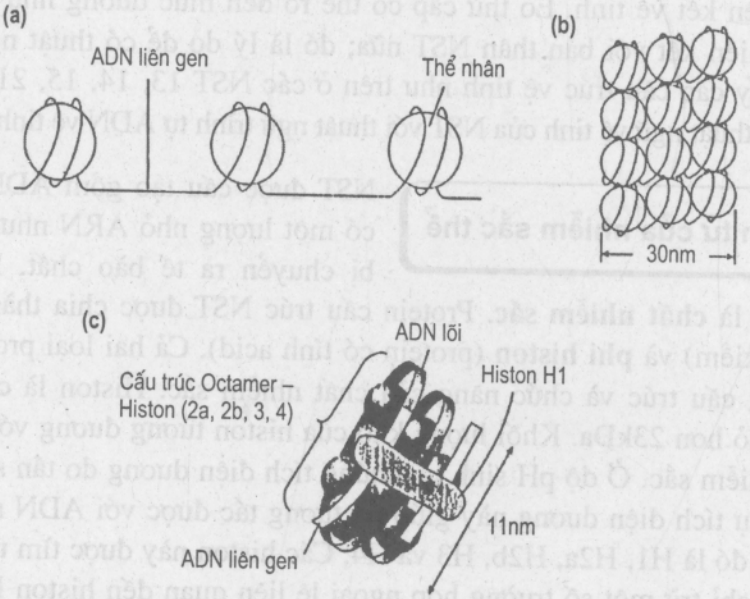
Vùng tổ chức hạch nhân thường được tìm thấy ở eo thứ cấp. Chúng gồm các trình tự mã hóa các gen rARN 5,8S; 18S và 28S được lặp lại kế tiếp nhau. Ở hầu hết các loài, các gen rARN 5S cụm lại tại một vùng trong hệ gen. Ở người, các vùng tổ chức hạch nhân thường được tìm thấy ở các cánh ngắn của tất cả các NST tâm lệch, ngoại trừ NST Y. Mỗi vùng tổ chức hạch nhân gồm khoảng 80 - 100 đoạn trình tự lặp lại. Trong kỳ trung gian, vùng tổ chức hạch nhân giãn xoắn và nhân con hình thành quanh nó; các vùng tổ chức hạch nhân của các NST khác nhau có thể phối hợp để hình thành một nhân con. Khi tế bào chuyển sang kỳ giữa của quá trình nguyên phân, các NST có thể vẫn còn gắn với nhau ở cánh ngắn. Hiện tượng đó được gọi là sự liên kết vệ tinh. Eo thứ cấp có thể rõ đến mức dường như vùng đầu mút của NST không còn liên kết với bản thân NST nữa; đó là lý do để có thuật ngữ vệ tinh NST. Ở người, có thể thấy các cấu trúc vệ tinh như trên ở các NST 13, 14, 15, 21 và 22 (*Hình 2.2*). Không nên nhầm thuật ngữ vệ tinh của NST với thuật ngữ trình tự ADN vệ tinh (xem mục II.D).

Cấu trúc phân tử của nhiễm sắc thể

NST được cấu tạo gồm ADN và protein; cũng có một lượng nhỏ ARN nhưng chỉ là để chuẩn bị chuyển ra tế bào chất. Hỗn hợp ADN và protein được gọi là **chất nhiễm sắc**. Protein cấu trúc NST được chia thành hai loại: **histon** (protein có tính kiềm) và **phi histon** (protein có tính acid). Cả hai loại protein đều có vai trò quan trọng trong cấu trúc và chức năng của chất nhiễm sắc. Histon là các protein có khối lượng phân tử nhỏ hơn 23kDa. Khối lượng khô của histon tương đương với ADN trong thành phần của chất nhiễm sắc. Ở độ pH sinh lý, chúng tích điện dương do tần số amino acid lysin và arginin cao. Sự tích điện dương này giúp nó tương tác được với ADN mang điện tích âm. Có 5 loại histon, đó là H1, H2a, H2b, H3 và H4. Các histon này được tìm thấy ở mọi loài sinh vật nhân chuẩn, chỉ trừ một số trường hợp ngoại lệ liên quan đến histon H1. Trình tự amino acid của các protein histon có tính bảo thủ rất cao trong suốt quá trình tiến hoá cho thấy các phân tử này có vai trò đặc biệt quan trọng, mang ý nghĩa sống còn đối với các sinh vật nhân chuẩn. Điều đó cũng được minh chứng bằng cấu trúc ổn định của các thể nhân trong cấu tạo chất nhiễm sắc khi so sánh giữa các loài khác nhau.

Về cấu trúc, các thể nhân gồm lõi histon bao quanh là vòng ADN. Phần lõi gồm hai đĩa song song, mỗi đĩa gồm bốn phân tử histon là H2a, H2b, H3 và H4. Phân tử ADN chạy quanh vành đĩa và gắn với một phân tử histon H1 nằm ngoài thể nhân. Vòng xoắn ADN cuốn quanh phần lõi thể nhân gồm 146 cặp base. Độ dài đoạn nối giữa các thể nhân có thể thay đổi khác nhau ở các loài khác nhau. Ở người, nó gồm khoảng 60 cặp base nên độ dài tổng cộng của

phân đoạn ADN tương ứng với một thể nhân là khoảng 200 cặp base. Đây là mức độ đóng gói cơ bản của ADN trong chất nhiễm sắc. Sự đóng gói tiếp theo phụ thuộc vào protein H1 (nằm ngoài lõi thể nhân). Các phân tử H1 có thể tương tác với nhau để cuộn các thể nhân thành cấu trúc cuộn xoắn có đường kính ~30nm. Đây chính là đường kính của các sợi nhiễm sắc thường nhìn thấy ở các ảnh chụp từ kính hiển vi điện tử. Tuy vậy, trong NST chất nhiễm sắc còn có những mức độ cuộn xoắn cao hơn. Cấu trúc của thể nhân được minh họa trên *Hình 2.3*. Các mức độ cuộn xoắn cao hơn được tìm thấy trong nhân tế bào. Mức độ cuộn xoắn cao nhất là các NST ở kỳ giữa của quá trình phân bào. Tổ chức của các cấu trúc này gồm các sợi nhiễm sắc gắn kết với phần khung NST. Phần khung này chủ yếu gồm protein có tính acid (phi histon) là **topoisomerase II.** Có những vùng đặc biệt trên phân tử ADN gồm vài trăm cặp base, giàu adenin và thymin, được gọi là các **vùng gắn kết của khung SAR** (*s*caffold *a*ttachment *r*egion) có vai trò liên kết trực tiếp phân tử ADN vào cấu trúc khung của NST. Vật liệu nằm giữa các vùng gắn kết khung này được sắp xếp như những chiếc thòng lọng có độ dài khác nhau. ADN sẽ lộ ra khi tiêu bản NST ở kỳ giữa được loại bỏ histon. Các ảnh hiển vi điện tử cho thấy các thòng lọng ADN dài toả ra xuất phát từ phần khung của NST. Các protein phi histon của nhân tham gia vào nhiều hoạt động điều hòa chức năng của chất nhiễm sắc, kể cả việc điều hoà sự hoạt động của các gen, phối hợp hoạt động với các yếu tố phiên mã vốn có vai trò chủ yếu.



Hình 2.3. **Thể nhân. (a) Sơ đồ một số thể nhân. (b) Các thể nhân xoắn thành cuộn đường kính 30nm. (c) Cấu trúc cơ bản của một thể nhân gồm ADN và protein histon.**

Chất nhiễm sắc hoạt động chức năng và không hoạt động chức năng

Không phải tất cả mọi vùng của chất nhiễm sắc đều tham gia hoạt động như nhau trong quá trình phiên mã. Một số vùng nhiễm sắc không tham gia vào quá trình phiên mã. Chúng được gọi là **chất dị nhiễm sắc** để phân biệt với vùng chất nhiễm sắc hoạt động, gọi là **chất nguyên nhiễm sắc**. Dưới kính hiển vi điện tử, chất dị nhiễm sắc gồm dày đặc các sợi nhiễm sắc. Có thể phân biệt chất dị nhiễm sắc với chất nguyên nhiễm sắc nhờ kính hiển vi quang học sau khi nhuộm màu NST. Một số vùng chất nhiễm sắc là dị nhiễm sắc ở tất cả các

mô và ở mọi giai đoạn phát triển cá thể. Những vùng này được gọi là **chất dị nhiễm sắc bắt buộc** (hay **dị nhiễm sắc cơ yếu**) và có thể phát hiện được nhờ kỹ thuật nhuộm băng C. Như đã nói ở trên, tâm động thường được phát hiện dễ dàng nhờ băng C. Có sự liên kết chặt chẽ giữa protein và các đoạn trình tự ADN vệ tinh ở vùng này giúp tránh sự mất đi của trình tự ADN trong quá trình nhuộm băng C. Vì vậy, vùng này nhuộm màu rất mạnh. Những vùng dị nhiễm sắc bắt buộc khác có thể kể đến bao gồm cánh dài của NST Y ở người và các vùng trên NST giới tính của các động vật khác. Băng C được dùng phổ biến hơn cho thực vật vì các NST thực vật thường khó tạo được băng G rõ nét.

Một số vùng của chất nhiễm sắc có thể tồn tại cả ở trạng thái nguyên nhiễm sắc hoặc dị nhiễm sắc, và được gọi là vùng chất **dị nhiễm sắc không bắt buộc (dị nhiễm sắc tùy ý)**. Động vật có vú giới cái thường có hai NST X, nhưng một chiếc thường không hoạt động. Nó được chuyển thành chất dị nhiễm sắc và được quan sát thấy như một chấm nhỏ đậm đặc bên cạnh nhân ở kỳ trung gian, gọi là **thể Barr** hay **chất nhiễm sắc X**. Bằng cách đó, dường như có sự bù trừ lượng chất nhiễm sắc giữa con đực và con cái, vì ở con đực chỉ có một NST X còn NST Y được cấu tạo chủ yếu từ chất dị nhiễm sắc bắt buộc. Sự bất hoạt một NST X trong cặp NST giới tính X ở giới cái là một sự kiện ngẫu nhiên. Có khoảng một nửa số tế bào của động vật có vú giới cái có NST X bắt nguồn từ bố bất hoạt, còn ở nửa kia là NST X bắt nguồn từ mẹ. Trong quá trình hình thành giao tử, NST X bất hoạt được hoạt hóa trở lại.

Sự biến đổi số lượng nhiễm sắc thể

Số lượng NST có thể thay đổi do những sai sót trong quá trình giảm phân. Những sai sót đó có thể dẫn đến việc những giao tử tạo thành có thêm hoặc mất đi một phần vật chất di truyền. Nếu các giao tử đó tồn tại được, chúng sẽ sinh ra đời con có các tế bào mang phần NST bất thường đó. Các sai sót xảy ra trong quá trình **nguyên phân** tạo ra những cá thể mang một số tế bào trong cơ thể với nhiều hơn một bộ NST bình thường. Các cá thể như vậy được gọi là **thể khảm.**

Đa bội thể là hiện tượng số lượng NST trong tế bào tăng hoặc giảm một số nguyên lần bộ NST đơn bội. Dãy đa bội nguyên mô tả các mức đa bội khác nhau. Chẳng hạn dãy đa bội nguyên ở người được mô tả trên *Bảng 2.1*. Mặc dù thể tam bội và tứ bội chiếm tới 10% số bào thai bị sảy thai ngẫu nhiên nhưng chưa bao giờ quan sát thấy ở giai đoạn trẻ sơ sinh. Hiện tượng đa bội thể xảy ra rất hiếm ở động vật nhưng xảy ra phổ biến ở thực vật và có vai trò quan trọng trong tiến hoá.

Hiện tượng dị đa bội là sự sai lệch nhỏ về số lượng NST so với dãy đa bội nguyên. Hiện tượng **thể không** là trường hợp vắng mặt hoàn toàn bản sao của bất kỳ một NST nào. Khi tế bào có một bản sao của một NST nào đó được gọi là **thể một**; có hai bản sao được gọi là **thể hai**; có ba bản sao được gọi là **thể ba**. Thể một thường thấy ở giao tử, thể hai thường thấy ở các tế bào soma. *Bảng 2.2* cho thấy một số ví dụ về các hội chứng gây ra do dị đa bội ở người.

Cũng có trường hợp dị đa bội cùng xảy ra trên nhiều NST. Song, trong phần lớn trường hợp các bào thai bị sảy thai sau đó. Có tổng cộng khoảng 4% các bào thai ở người là thể ba. Những trường hợp dị đa bội sống sót được thường chết sớm. Còn những cá thể có một phần của một NST nào đó được nhân lên ba lần thì biểu hiện ít nghiêm trọng hơn. Tuy vậy, các dạng thể ba xuất hiện ở người được nêu ở trên chỉ đúng đối với các NST nhỏ. Nếu có sự dư thừa một bản sao của một NST lớn thì hậu quả cũng lớn vì có thể liên quan đến rất nhiều gen trong quá trình phát triển bào thai.

Bảng 2.1. **Dãy đa bội nguyên ở người**

Số lượng NST	Mức bội thể	Nơi thường tìm thấy
23	Đơn bội	Giao tử
46	Lưỡng bội	Tế bào soma
69	Tam bội	
92	Tứ bội	

Bảng 2.2. **Một số thể dị đa bội ở người**

Hội chứng	NST	Tần số tương đối
Thừa một NST thường		
Hội chứng Edward	18	1/5000
Hội chứng Pateau	13	1/5000
Hội chứng Down	21	1/750
Dị đa bội ở NST giới tính		
Hội chứng Turner	XO	1/10 000
Hội chứng Klinfelter	XXY	1/2000
Hội chứng ba X	XXX	1/2000

Hiện tượng dị đa bội ở NST giới tính khá phổ biến, mặc dù NST X lớn và chứa nhiều gen. Điều đó là do: trong bất kỳ tế bào nào của động vật có vú chỉ có một NST X được hoạt hóa, vì vậy sự có mặt của những bản sao thừa ít ảnh hưởng đến sức sống của tế bào. Hội chứng Turner (XO) hiếm gặp hơn các thể dị đa bội khác, nó thường thấy nhiều hơn ở dạng thể khảm, lẫn lộn với các tế bào (XX) nhờ vậy nó có thể sống sót.

Giao tử của thể mang	Giao tử bình thường	
21, 14	21, 14	Sống, bình thường
21/14	21, 14	Thể mang, sống
21, 21/14	21, 14	Thể ba NST 21
14	21, 14	Thể không, chết
14, 21/14	21, 14	Không sống được
21	21, 14	Thể không, chết

***Hình 2.4.* Sự di truyền thể ba NST 21.** *Nếu một người là thể ba NST 21 (Hội chứng Down) do chuyển đoạn giữa NST 21 và NST 14 thì người đó sẽ sinh ra nhiều loại giao tử khác nhau bởi sự phân ly khác thường của NST chuyển đoạn ở giảm phân. Các giao tử này thụ tinh với các giao tử bình thường. Hình trên biểu diễn các giao tử do thể mang chuyển đoạn tạo ra và các giao tử bình thường cũng như kết quả sự kết hợp của các giao tử đó. Chuyển đoạn được ký hiệu là 21/14. Chú ý rằng các giao tử được sinh ra với tần số khác nhau.*

Các NST dị đa bội có thể xuất hiện do những sai sót ở cả lần phân bào thứ nhất hoặc thứ hai của quá trình giảm phân do sự không phân ly của các nhiễm sắc tử (NST con). Sự không phân ly đó có thể xảy ra do hai NST tương đồng không tách nhau ra ở giảm phân I, hoặc do tâm động không phân ly ở kỳ giữa của giảm phân II. Sự không phân ly NST do tâm động không phân cắt còn có thể xảy ra trong nguyên phân. Người ta cho rằng sự không phân ly NST xuất hiện như một sự kiện ngẫu nhiên giống như đột biến. Tuy vậy, chúng thường bị tác động bởi các yếu tố bên ngoài. Có thể thấy rõ điều đó ở thể ba NST 21 (cơ chế chính gây hội chứng Down). Tần số hội chứng Down tăng dần cùng với tuổi của mẹ. Trong một số trường hợp, hội chứng này có khuynh hướng theo dòng họ, do những người này có các thể mang là kết quả của sự chuyển đoạn NST 21. Trong quá trình sinh sản, những người này có thể truyền cả bản sao NST 21 bình thường và cả NST 21 chuyển đoạn sang cho con. Như vậy, đứa trẻ được truyền sẽ có ba bản sao một phần NST 21. Sự kiện đó được mô tả trên *Hình 2.4*.

II.B SỰ PHÂN CHIA TẾ BÀO (NGUYÊN PHÂN)

Chu trình tế bào ở sinh vật nhân chuẩn

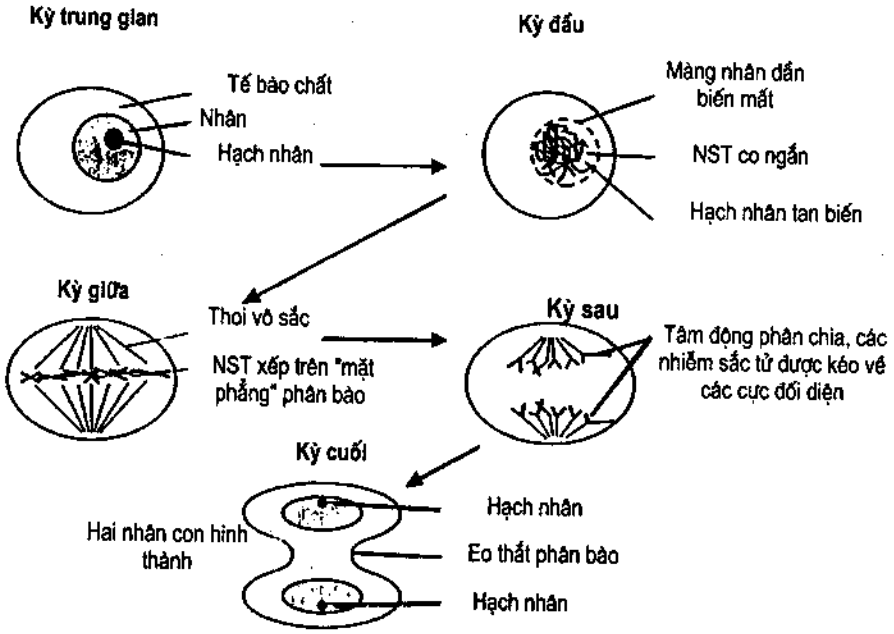
Tiến trình từ lần phân bào này đến lần phân bào sau là quá trình có tính chu kỳ, gọi là **chu trình tế bào**. Trong thời gian đó, tế bào phải sao chép các thành phần của tế bào và tổ chức phân phối các thành phần đó đồng đều giữa hai tế bào con. Ngoại trừ quá trình hình thành giao tử, nhân của các tế bào ở sinh vật nhân chuẩn phân chia theo hình thức **nguyên phân.** Cùng với quá trình đó, tế bào chất cũng phân chia nhờ **quá trình phân chia tế bào chất.** Những quá trình này có thể dễ dàng quan sát được ở các tế bào sống hoặc đã được cố định. Giai đoạn giữa hai lần phân chia tế bào liên tiếp (không kể pha phân bào) được gọi là **kỳ trung gian.** Sự sao chép ADN được thực hiện ở pha tổng hợp hay pha S trong kỳ trung gian. Trong kỳ trung gian, giai đoạn trước pha S được gọi là G1 (gap 1) và giai đoạn giữa pha S và pha phân bào tiếp theo được gọi là G2 (gap 2).

Nguyên phân và sự phân chia tế bào chất

Nguyên phân có nhiều điểm giống với giảm phân (là cơ chế phân chia tế bào giảm nhiễm để tạo thành giao tử). Do vậy, cần chú ý phân biệt để tránh nhầm lẫn giữa hai kiểu phân bào này. Các kỳ của nguyên phân được mô tả trên *Hình 2.5*.

Như đã nói ở trên, các pha G1, G2 và S hợp thành kỳ trung gian. Khi tế bào kết thúc pha G2, nó bước vào **kỳ đầu** của giảm phân. Trong kỳ này, từng nhiễm sắc thể hiện rõ dần vì các chất nhiễm sắc bắt đầu kết đặc. Lúc này, chúng có cấu trúc kép gồm hai **nhiễm sắc tử** vì ADN đã được sao chép trước đó từ pha S. Đến cuối kỳ đầu, màng nhân và nhân con bắt đầu phân hủy. Các protein nhân được giải phóng ra liên kết vào bề mặt của các nhiễm sắc thể đang kết đặc. Tế bào chuyển sang **kỳ giữa.** Vào giai đoạn này, màng nhân và nhân con không còn nữa và các nhiễm sắc thể được sắp xếp trên mặt phẳng xích đạo của tế bào. Đó là do hoạt động của các sợi thoi vô sắc có chứa protein tublin đi từ các cực phân chia của tế bào và gắn với tâm động của nhiễm sắc thể. Kỳ tiếp theo là **kỳ sau**, bắt đầu ngay khi tâm động phân cắt để phân tách các nhiễm sắc tử. Mỗi nhiễm sắc tử, lúc này nên gọi là nhiễm sắc thể con, được kéo về

hai cực đối diện của tế bào nhờ các sợi thoi phân bào. Khi hai bộ nhiễm sắc thể đã hoàn toàn tách về hai cực, chúng bắt đầu giải kết đặc, màng nhân hình thành quanh chất nhiễm sắc đang giải kết đặc và tế bào tạm thời có hai nhân; đây là giai đoạn **kỳ cuối**. Ở tế bào động vật, vào giai đoạn kỳ cuối, **rãnh phân cắt** hình thành qua vùng xích đạo của tế bào. Cấu trúc phân tách được cấu tạo từ các vi sợi chứa actin này thắt chặt lại nhanh chóng và cuối cùng chia tế bào thành hai tế bào con; đó là quá trình **phân chia tế bào chất**. Ở các tế bào thực vật, tế bào chất phân chia bằng cách hình thành một thành tế bào mới qua vùng xích đạo của tế bào.



Hình 2.5. **Sơ đồ quá trình nguyên phân ở tế bào động vật**

Sự điều hòa chu trình tế bào

Khái niệm này đề cập đến tốc độ các tế bào trải qua các giai đoạn của chu trình tế bào. Điểm quan trọng cần nắm vững là: ở các sinh vật đa bào, phần lớn các tế bào không phân chia đồng đều. Các tế bào không hoạt động, được coi là đang ở pha G_0, cần được kích thích trước khi chúng phân bào. Chỉ có các tế bào ở các mô đặc biệt như tế bào tủy xương của động vật và đầu rễ của thực vật là có tỷ lệ các tế bào phân chia cao. Các tế bào có thể được kích thích để thoát khỏi pha G_0 nhờ các hormon sinh trưởng đặc hiệu. Sau đó, chúng vượt qua pha G1. Giai đoạn này dài ngắn khác nhau tùy từng loại tế bào, nhưng một khi tế bào đã vượt qua điểm đặc biệt được gọi là **điểm giới hạn** hay **điểm R**, tế bào có xu hướng phải hoàn thành chu trình tế bào, nghĩa là tế bào tiến tới lần phân bào tiếp theo, rồi trở lại pha G1. Các pha khác của chu trình, gồm các pha S, G2 và phân bào có độ dài thời gian tương đối ổn định.

Tiến trình của chu trình tế bào là sự biểu hiện của tế bào nhằm đáp ứng lại hiệu ứng của các hormon kích thích sinh trưởng được điều hoà thông qua một quá trình gọi là **quá trình truyền tín hiệu**. Quá trình đó có sự tham gia của hàng loạt các phân tử kích ứng làm nhiệm vụ khuếch đại và truyền tín hiệu được khởi đầu bởi hormon sinh trưởng; cuối cùng thúc đẩy sự phiên mã của các gen tham gia vào quá trình sinh trưởng và phân chia của tế bào. Đồng thời, tiến trình của chu trình tế bào còn được điều hòa âm tính bằng hàng loạt các điểm kiểm tra tế

bào mà ở đó tế bào không thể vượt qua nếu không đáp ứng được những tiêu chí nhất định. Nhờ vậy, quá trình sinh trưởng và phân chia tế bào là một quá trình được điều hòa chặt chẽ vì các quá trình đó là thiết yếu cho sự phát triển của sinh vật đa bào. Những nhân tố chính quy định tiến trình của chu trình tế bào như vậy gồm nhóm enzym kinase, là những enzym làm nhiệm vụ bổ sung nhóm phosphat vào cơ chất của chúng. Sự phosphoryl hoá các protein đích đặc hiệu ở những giai đoạn khác nhau của chu trình tế bào là điều kiện cần thiết để chu trình tế bào có thể tiến triển. Các enzym kinase còn liên kết với nhóm protein thứ hai, gọi là các protein cyclin. Việc liên kết với các protein này giúp điều hoà hoạt tính của các enzym kinase. Khi không có cyclin, enzym kinase không hoạt động, và vì thế chúng được gọi là enzym kinase phụ thuộc cyclin. Ở những thời điểm khác nhau của chu trình tế bào, có các loại protein cyclin khác nhau.

Người ta đã xác định được nhiều điểm kiểm tra tế bào, trong đó có ít nhất ba điểm kiểm tra tương ứng với ba pha của chu trình tế bào, đó là ranh giới giữa pha G1 và pha S, giữa pha G2 và nguyên phân và một điểm ngay trong quá trình nguyên phân. Những điểm kiểm tra đó đảm bảo cho quá trình sinh trưởng và phân chia của tế bào diễn ra chính xác. Chẳng hạn, những tế bào mang những tổn thương ADN không được sửa chữa sẽ bị ngăn cản không chuyển sang được pha S; vì nếu không, các tổn thương trên ADN có thể sinh ra đột biến. Các tế bào không hoàn thành quá trình sao chép ADN sẽ không chuyển được sang được pha nguyên phân. Trong quá trình nguyên phân, tế bào không được phân chia nếu không có sự hình thành các sợi thoi vô sắc, hoặc khi các nhiễm sắc thể không gắn được vào nó.

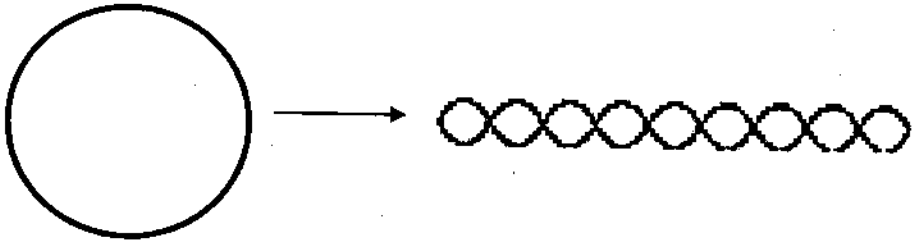
Các đột biến ở các gen tham gia vào con đường truyền tín hiệu điều hoà chu trình tế bào có thể là tác nhân gây ung thư. Những đột biến này được gọi là các gen gây ung thư (oncogene). Các gen tham gia vào các điểm kiểm tra nếu bị đột biến cũng thường gây nên sự hình thành khối u. Ở trạng thái bình thường, chúng được gọi là các gen ức chế khối u.

II.C HỆ GEN SINH VẬT NHÂN SƠ

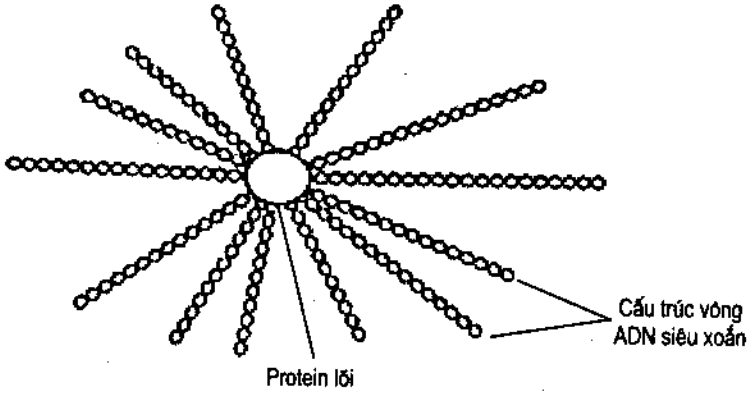
Nhiễm sắc thể sinh vật nhân sơ

Sinh vật được phân loại thành các dạng sơ khai như vi khuẩn được gọi là các **sinh vật nhân sơ** (prokaryote) và các sinh vật bậc cao được gọi là các **sinh vật nhân chuẩn** (eukaryote). Cấu trúc và tổ chức tế bào của các sinh vật nhân sơ và nhân chuẩn khác nhau. Các tế bào của sinh vật nhân chuẩn có cấu trúc bên trong phức tạp với các bào quan có màng bao bọc và nhân riêng biệt. Các tế bào ở sinh vật nhân sơ không có tổ chức như vậy và không có nhân riêng biệt. Các sinh vật nhân sơ được chia thành **vi khuẩn thật,** trong đó có *E. coli* và một nhóm vi khuẩn không bình thường. Khác biệt với vi khuẩn thật là **vi khuẩn cổ**. Hầu hết các hiểu biết về vi khuẩn có được đều dựa trên *E. coli* và các vi khuẩn thật khác. Sinh vật nhân sơ có một phân tử ADN dạng vòng duy nhất được gọi là **nhiễm sắc thể vi khuẩn** chứa gần như toàn bộ các gen chính yếu của vi khuẩn. Phân tử ADN này rất dài so với kích thước của tế bào vi khuẩn. Để xếp gọn được trong tế bào, phân tử ADN này phải được nén lại nhờ **quá trình siêu xoắn**. Các enzym gọi là **topoisomerase** tạo ra các vòng uốn bổ sung để cuộn phân tử ADN trở nên chặt hơn (*Hình 2.6*). Topoisomerase tác động bằng cách bẻ

gãy chuỗi polynucleotid và quay hai đầu phân tử ADN lại với nhau. Sau đó enzym nối các đầu lại với nhau và như vậy chuỗi polynucleotid cuộn lại. Quá trình này được gọi là sự siêu xoắn dương tính. Topoisomerase còn có thể cởi bỏ xoắn trong quá trình ngược lại được gọi là siêu xoắn âm tính bằng cách tạo ra các vòng uốn theo hướng ngược lại. Sự siêu xoắn cũng xảy ra ở các tế bào nhân chuẩn để bao gói ADN thành các nhiễm sắc thể nhân chuẩn.



Hình 2.6. **Sự cuộn xoắn của phân tử ADN dạng vòng**



Hình 2.7. **Cấu trúc nucleoid ở *E. coli***

Mặc dù tế bào vi khuẩn không có nhân riêng biệt nhưng các phân tích bằng kính hiển vi điện tử cho thấy tế bào vi khuẩn có một vùng trung tâm sẫm màu chứa ADN và protein được gọi là vùng nhân (nucleoid) và vùng bao quanh được gọi là tế bào chất. Cấu trúc chi tiết của vùng nhân chưa được hiểu biết và mô tả đầy đủ, song một số dẫn liệu cho thấy vùng nhân có lõi protein trung tâm, từ đó các thòng lọng ADN dạng siêu xoắn toả ra (*Hình 2.7*). Một số protein tách chiết từ vùng nhân giống với protein histon ở các nhiễm sắc thể sinh vật nhân chuẩn. Do vậy, người ta cho rằng các protein này có vai trò cuộn xoắn phân tử ADN thành cấu trúc chặt hơn. Độ dài của nhiễm sắc thể nhân sơ so với kích thước của tế bào cho thấy sự sao chép và phân chia các phân tử ADN trong quá trình phân bào là một quá trình khá phức tạp. Quá trình đó chỉ có thể thực hiện được bằng cách phân tử ADN có những điểm gắn riêng biệt trên màng tế bào để từ đó các phân tử ADN đã sao chép có thể tách ra khi tế bào phân chia.

Tổ chức hệ gen sinh vật nhân sơ

Hầu hết các gen của vi khuẩn đều nằm trên nhiễm sắc thể của vi khuẩn. Nhưng có một số gen nằm trên phân tử ADN dạng vòng nhỏ được gọi là **plasmid** bên ngoài ADN nhiễm sắc thể của vi khuẩn. Để minh họa cho đặc điểm tổ chức của các gen trên nhiễm sắc thể vi khuẩn, xin lấy ví dụ ở *E. coli*. Vi khuẩn này có 2800 gen do một phân tử ADN mang 4,6 triệu cặp base mã hóa. Toàn bộ trình tự các base trên phân tử ADN nhiễm sắc thể của *E. coli* và vị trí tương ứng của chúng trên nhiễm sắc thể đã được xác định.

Một số gen được sắp xếp thành các **operon** mã hoá cho các protein có chức năng có liên quan với nhau và cùng được điều hoà biểu hiện đồng thời. Những gen khác nằm ở các vị trí ngẫu nhiên trên nhiễm sắc thể, không có tổ chức rõ ràng. Hầu hết các gen chỉ có một bản sao duy nhất, ngoại trừ các gen mã hóa rARN tồn tại thành cụm gồm bảy bản sao. Cách tổ chức các gen tương tự như vậy cũng gặp ở các vi khuẩn thật khác và bản đồ gen rất giống nhau ở những loài có quan hệ họ hàng gần gũi. Khoảng 75% trình tự ADN của nhiễm sắc thể vi khuẩn là trình tự của các gen. 25% còn lại là **trình tự ADN liên gen** nằm ở vùng ngăn cách giữa các gen. Một số đoạn trình tự ADN liên gen có những chức năng quan trọng như là **điểm khởi đầu phiên mã**. Những đoạn khác có thể có vai trò tương tác với các protein bao gói ADN. Kích thước hệ gen sinh vật nhân sơ khác nhau giữa các loài vi khuẩn khác nhau. Sự khác nhau về kích thước phản ánh sự khác nhau về số lượng gen và cả độ dài những đoạn ADN liên gen.

Plasmid

Đó là phân tử ADN dạng vòng bổ sung nằm ngoài nhiễm sắc thể có mặt ở hầu hết các loài vi khuẩn. Plasmid mang các gen không có trên nhiễm sắc thể vi khuẩn và thường quy định những đặc điểm có ích cho vi khuẩn, chẳng hạn như tính kháng các chất kháng sinh. Plasmid có điểm khởi đầu sao chép riêng và vì thế nó có thể sao chép độc lập với ADN nhiễm sắc thể. Một số plasmid nhỏ sử dụng enzym của tế bào để sao chép. Những plasmid khác lồng ghép vào hệ gen và cùng sao chép với nhiễm sắc thể vi khuẩn. Dạng lồng ghép này của plasmid được gọi là **episom** và có thể trải qua nhiều lần phân bào trước khi trở lại là một plasmid độc lập.

Có nhiều loại plasmid được tìm thấy ở vi khuẩn và một loài vi khuẩn có thể chứa một số loại plasmid khác nhau. Có thể phân loại plasmid theo các gen mà chúng mang và đặc điểm của các gen plasmid xác định cho tế bào vật chủ. Về cơ bản, có thể phân biệt năm loại plasmid chính như sau:

– **Plasmid kháng** (plasmid R). Loại plasmid này mang các gen làm cho vi khuẩn kháng với các chất kháng sinh như ampicillin và chloramphenicol. Cơ chế tạo ra tính kháng có thể khác nhau. Chẳng hạn, plasmid RP4 tìm thấy ở *Pseudomonas* và một số vi khuẩn khác. Việc truyền plasmid R giữa các loại vi khuẩn khác nhau là một cơ chế phổ biến tính kháng với các chất kháng sinh giữa các loài vi khuẩn khác nhau và cần được quan tâm khi xử lý quá trình lây nhiễm ở vi khuẩn.

– **Plasmid sinh sản** (Plasmid F, *F = fertility*). Loại plasmid này cho phép các gen được truyền giữa các tế bào vi khuẩn trong **quá trình tiếp hợp**. Plasmid F chứa các gen điều khiển việc truyền plasmid F từ tế bào vi khuẩn này sang tế bào vi khuẩn khác qua một cấu trúc hình ống được gọi là **lông giới tính**. Trong quá trình tiếp hợp, plasmid F có thể mang một số gen khác có nguồn gốc từ nhiễm sắc thể và truyền cho thể nhận.

– **Plasmid col**. Loại plasmid này mang các gen mã hoá cho protein colicin có thể giết chết các vi khuẩn khác. Ví dụ như ColE1 của *E. coli*.

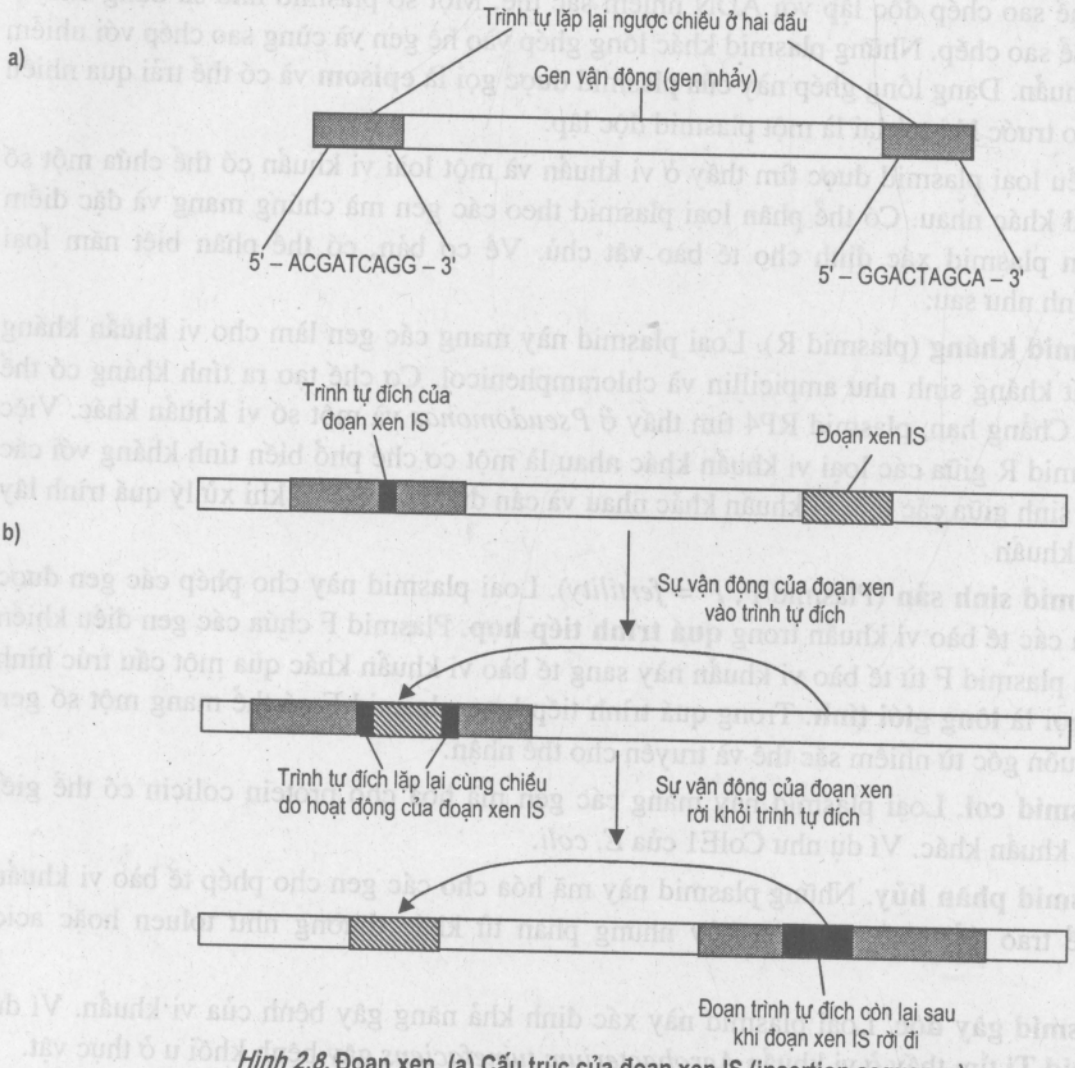
– **Plasmid phân hủy**. Những plasmid này mã hóa cho các gen cho phép tế bào vi khuẩn chủ có thể trao đổi chất và phân hủy những phân tử khác thường như toluen hoặc acid salicylic.

– **Plasmid gây độc**. Loại plasmid này xác định khả năng gây bệnh của vi khuẩn. Ví dụ như **plasmid Ti** tìm thấy ở vi khuẩn *Agrobacterium tumefaciens* gây bệnh khối u ở thực vật.

Kích thước các loại plasmid có thể rất khác nhau, nhỏ nhất khoảng 1kb và lớn nhất tới 250kb. Mỗi loại plasmid thích ứng với những loại vi khuẩn chủ khác nhau. Một số plasmid có mặt ở nhiều loài vi khuẩn, một số khác chỉ có mặt ở một vài loài. Số lượng plasmid trong một tế bào vi khuẩn cũng có thể khác nhau. Những plasmid chỉ tồn tại với một hay hai bản sao trong tế bào được gọi là các **plasmid hiếm**. Một số loại plasmid khác có số bản sao cao hơn, có thể là 10 hoặc nhiều hơn trong tế bào, được gọi là các **plasmid dư thừa**. Có sự hạn chế về sự tồn tại đồng thời của nhiều loại plasmid trong cùng một tế bào. Đặc tính này được gọi là **tính không tương hợp của các plasmid**. Các plasmid cùng tồn tại được trong một tế bào vi khuẩn phải thuộc các nhóm không tương hợp khác nhau.

Gen nhảy (transposon) ở vi khuẩn

Đó là các đoạn ADN có khả năng vận động trong hệ gen. Các "gen nhảy" có cả ở sinh vật nhân sơ và sinh vật nhân chuẩn. Ở sinh vật nhân sơ, chúng xuất hiện cả trên nhiễm sắc thể và trên plasmid. Sự chuyển vị trí của gen nhảy gọi là **sự vận động của gen nhảy** phụ thuộc vào sự tái tổ hợp giữa các đoạn trình tự ADN. Các gen nhảy điển hình là những đơn vị "tự trị" và mã hóa cho enzym transposase xúc tác cho sự vận động của chính nó.



Hình 2.8. **Đoạn xen. (a) Cấu trúc của đoạn xen IS (insertion sequence). (b) Đột biến gen do sự chuyển vị trí (vận động) của đoạn xen**

Đến nay đã phát hiện ra nhiều loại gen nhảy khác nhau. Loại đầu tiên là các **đoạn xen** tìm thấy ở *E. coli* (*Hình 2.8a*). Đã xác định được một số kiểu đoạn xen khác nhau, trong đó có những đoạn xen có đến 10 bản sao trong cùng hệ gen vi khuẩn. Sự vận động của gen nhảy nhìn chung xảy ra với tần số thấp, chỉ vào khoảng $1/10^3$ - $1/10^4$ lần phân bào. Các gen nhảy có thể được truyền giữa các vi khuẩn trong quá trình tiếp hợp và cũng có thể được truyền giữa các loài có quan hệ họ hàng. Điểm đặc trưng của các đoạn xen là chúng có đoạn trình tự lặp lại ngược chiều ngắn ở cả hai đầu. Đó là những đoạn trình tự lặp lại trong đó hai bản sao định hướng ngược chiều nhau. Khi đoạn xen vận động thì trình tự ADN tại vị trí đích được sao chép lặp đoạn sao cho vùng biên ở hai đầu đoạn xen mang hai đoạn **trình tự lặp lại cùng chiều**. Khả năng các gen nhảy vận động được trong hệ gen đồng nghĩa với việc chúng có thể gây đột biến ở gen đích khi chúng xen vào gen đích. Khi chúng vận động sang một ví trí khác thì đoạn lặp trình tự lặp lại cùng chiều ngắn vẫn còn lại. Vì vậy, gen vẫn bị đột biến ngay cả khi gen nhảy đã rời đi.

Vi khuẩn cổ

Đây là nhóm sinh vật khác với vi khuẩn thật. Ngày nay người ta thấy rằng nhóm vi khuẩn này khác cả với các sinh vật nhân chuẩn và các sinh vật nhân sơ khác. Vi khuẩn cổ gồm ba nhóm có quan hệ họ hàng với nhau: các **vi khuẩn sinh khí metan**, các **vi khuẩn ưa nhiệt cực đoan** và các **vi khuẩn ưa mặn cực đoan**. Phần lớn các vi khuẩn cổ sống trong các môi trường cực đoan nên rất khó nuôi cấy chúng và vì thế khó nghiên cứu. Về mặt hóa sinh, chúng khác với các vi khuẩn thật về cấu tạo và cấu trúc thành tế bào. Về mặt di truyền, có một số khác biệt liên quan đến các gen mã hóa rARN như sự có mặt của các intron, sự khác biệt về cấu trúc và tổ chức gen.

II.D HỆ GEN NGƯỜI

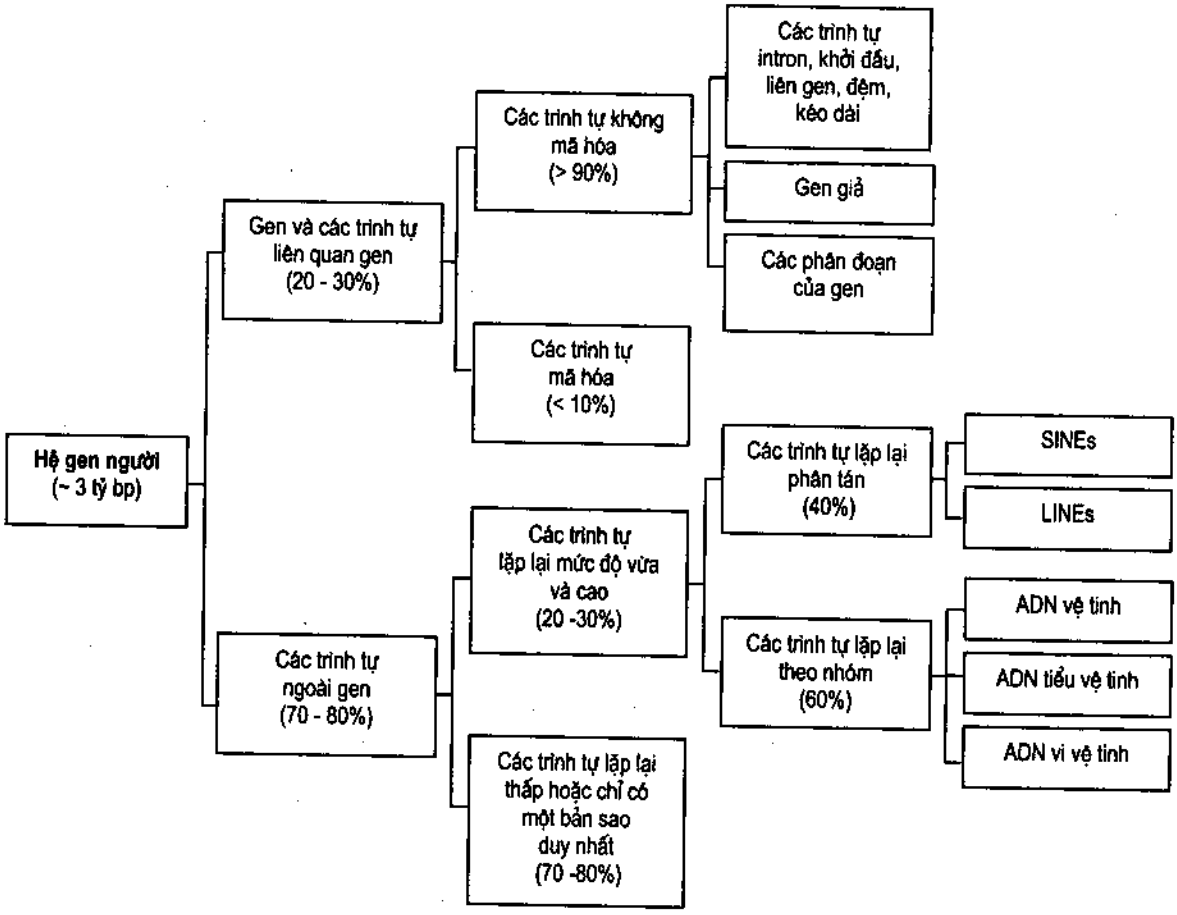
Hệ gen người

Thuật ngữ này được dùng để mô tả các kiểu trình tự khác nhau, hợp lại để tạo nên ADN trong tế bào người. ADN trong hệ gen đơn bội ở người dài hơn ba tỷ cặp base (bp) và ước tính chứa khoảng 40.000 gen. ADN được phân bố trên 23 nhiễm sắc thể, mỗi nhiễm sắc thể là một phân tử ADN sợi kép dài có kích thước khoảng 55 - 250 triệu bp. Các gen và các trình tự liên quan đến gen chiếm khoảng 25% toàn bộ hệ gen (*Hình 2.9*). Phần còn lại được gọi là ADN liên gen có chức năng sinh học chưa biết đầy đủ.

Gen ở người

Thông tin mã hoá trong gen biểu hiện qua chuỗi các đoạn trình tự ADN được gọi là các exon phân cách nhau bởi các đoạn trình tự không mã hoá được gọi là các intron. Các gen có kích thước cũng như số lượng các exon và kích thước các intron rất khác nhau. Một số gen như gen histon H4 chỉ dài vài trăm cặp base. Các gen khác như gen mã hóa yếu tố VIII dài vài trăm kilobase (kbp) và chứa rất nhiều intron kích thước lớn tới mức đoạn trình tự mã hoá thực sự chỉ chiếm vài phần trăm toàn bộ trình tự của gen. Có nhiều đoạn trình tự bổ sung liên quan đến gen. Đoạn trình tự dẫn đầu và theo sau ở đầu 5' và 3' được phiên mã song không được dịch mã. Đoạn trình tự khởi

động nằm ngược dòng với điểm khởi đầu phiên mã và điều hòa việc tổng hợp mARN từ gen. Đoạn khởi động có thể dài tới 1kbp, còn các đoạn trình tự điều hoà có ảnh hưởng đến sự phiên mã có thể nằm ở những vị trí xa hơn nữa.



Hình 2.9. **Các loại trình tự ADN trong hệ gen người**

Họ gen

Một số gen tồn tại thành nhiều bản sao (copy) có trình tự giống hệt nhau hoặc có liên quan với nhau để có thể nhóm thành họ. Các họ gen có thể được tổ chức theo nhiều cách, như: (i) tất cả các gen trong họ gen đều nằm trên cùng một locut gen. Ví dụ như họ gen hormon sinh trưởng có năm gen cụm lại trên nhiễm sắc thể 17; (ii) các gen thuộc cùng một họ gen có thể nằm trên các locut khác nhau. Ví dụ, năm gen của họ gen aldolase nằm trên các nhiễm sắc thể khác nhau; (iii) các gen của một họ gen có thể tồn tại thành nhiều cụm trên các nhiễm sắc thể khác nhau. Ví dụ các gen homeobox tồn tại thành bốn cụm trên các nhiễm sắc thể khác nhau, mỗi cụm chứa khoảng 10 gen. Ở một số họ gen, tất cả các gen đều giống hệt nhau và cùng mã hoá một protein mà tế bào cần với số lượng lớn như các protein histon. Ở những họ gen khác, các gen không giống hệt nhau mà có những trình tự khác biệt. Trong một số trường hợp, sự khác biệt lớn tới mức chúng mã hoá các protein tuy có quan hệ với nhau nhưng có những đặc tính khác biệt hẳn. Ví dụ như họ gen α- và β-globin gồm các gen biểu hiện ở các giai đoạn khác nhau của quá trình phát triển phôi và ở người trưởng thành.

Gen giả

Trong hệ gen người, có sự tồn tại một số lượng lớn các thành viên của các họ gen mang một hoặc một số đột biến làm mất chức năng, vì thế chúng không còn hoạt động chức năng nữa và không sinh ra các protein có chức năng sinh học. Gen giả đơn giản là phiên bản bị đột biến của gen bố, mẹ. Thường các đột biến này là các đột biến vô nghĩa, sinh ra các bộ ba kết thúc chuỗi làm kết thúc sớm sự dịch mã. Có liên quan tới các gen giả là các **gen giả gia công**, đó là các bản sao ADN của phân tử mARN. Các gen giả gia công không có intron và vì chúng không có đoạn trình tự khởi động nên không phiên mã được và không tạo ra protein. Chưa có bằng chứng khẳng định bằng cách nào các gen giả gia công xuất hiện trong hệ gen người, tuy nhiên có thể do quá trình phiên mã ngược mARN thành ADN sợi kép rồi sau đó được lồng ghép vào hệ gen. Một nhóm gen không hoạt động khác là các **đoạn gen**, chúng không có vùng đầu 5' và 3' của gen bố, mẹ. Người ta cho rằng chúng xuất hiện do mất đoạn hoặc do tái tổ hợp các đoạn phân cắt của các gen tiền thân.

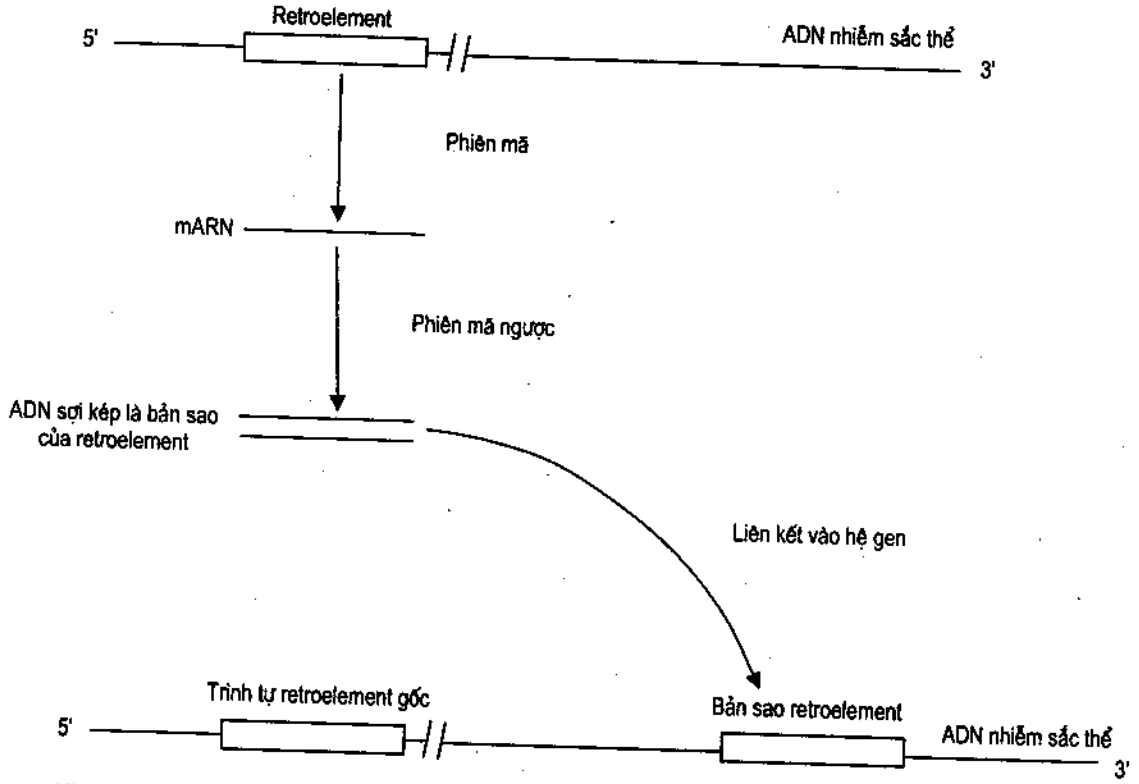
ADN liên gen

Phần ADN này của hệ gen người bao gồm các đoạn trình tự tồn tại ngoài bản thân các gen và các đoạn trình tự liên quan đến gen như đã mô tả ở trên (*Hình 2.9*). ADN ngoài gen bao gồm các đoạn trình tự không phải là một phần của gen (như exon và intron), không phải là phần có liên quan đến gen (đoạn dẫn đầu, đoạn theo sau, đoạn khởi động và các yếu tố điều hòa ở xa), và cũng không phải là các gen giả và các đoạn gen. Mặc dù các trình tự ngoài gen chiếm hầu hết ADN của hệ gen người (70 - 80%) song chức năng của chúng chưa được rõ. Hầu hết các đoạn trình tự ADN ngoài gen đều là những trình tự duy nhất hoặc tồn tại dưới dạng số ít bản sao. Phần còn lại (20 - 30%) là những đoạn trình tự lặp lại ở mức trung bình hoặc cao có thể phân tán suốt hệ gen hoặc sắp xếp kế tiếp nhau thành một chuỗi dài một cách ngẫu nhiên.

Các đoạn trình tự lặp lại phân tán

Có hai loại trình tự lặp lại phân tán là các yếu tố hệ gen nhân ngắn phân tán SINE (*s*hort *i*nterspersed *n*uclear *e*lement) và các yếu tố hệ gen nhân dài phân tán LINE (*l*ong *i*nterspersed *n*uclear *e*lement). Ví dụ rõ nhất của SINE là các yếu tố *Alu*. Các đoạn trình tự này không giống hệt nhau nhưng chúng đủ giống nhau để được phân loại thành một họ. Chúng có độ dài trung bình 250bp và có khoảng 700.000 bản sao nằm rải rác trong hệ gen, cả ở trong intron của một số gen. Nguồn gốc của các yếu tố *Alu* còn chưa rõ, nhưng người ta cho rằng chúng xuất hiện từ một hoặc một số gen giả gia công. Trong đó, ở một giai đoạn nào đó các đoạn trình tự này có được khả năng vận động, cho phép chúng sao chép và vận động tới các vị trí khác nhau trong hệ gen (vì vậy chúng còn được gọi là "gen nhảy").

LINE cũng giống như SINE nhưng có trình tự dài hơn. Ví dụ rõ nhất là LINE *L1* dài khoảng 6500bp, tồn tại với khoảng 60.000 bản sao trong hệ gen. LINE là một loại yếu tố di truyền vận động theo kiểu phiên mã ngược (retroelement, retrotransposon). Nghĩa là, chúng là các đoạn trình tự có khả năng vận động trong hệ gen thông qua một quá trình vận động mà trong đó một bản sao của chúng được tạo ra bởi cơ chế phiên mã ngược (*Hình 2.10*). Hầu hết các yếu tố *L1* bị cắt xén, tuy nhiên để quá trình vận động có thể xảy ra, cần một bản sao có độ dài đầy đủ bao gồm cả gen mã hóa cho enzym phiên mã ngược.



Hình 2.10 **Sự vận động của yếu tố di truyền vận động kiểu phiên mã ngược (retroelement)**

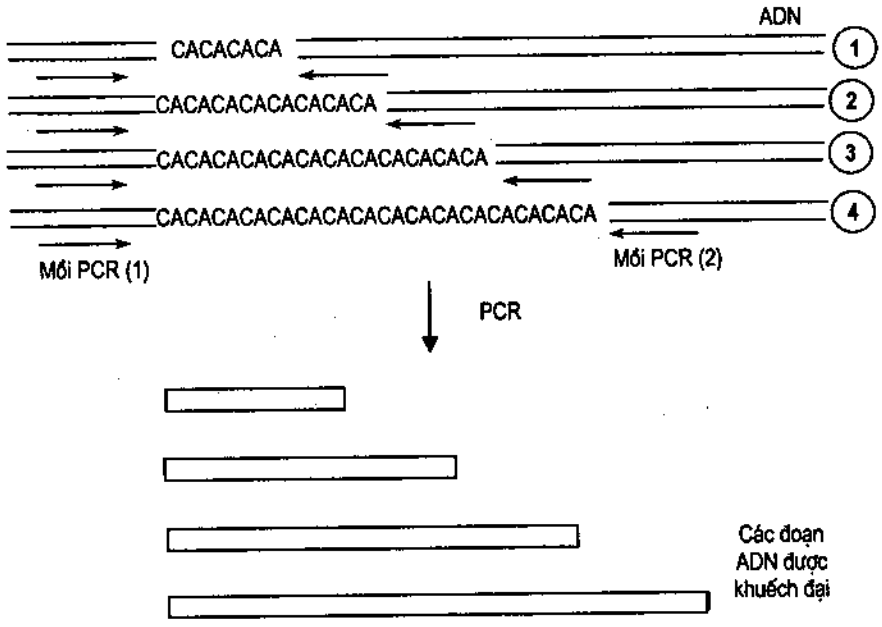
Các đoạn trình tự lặp lại co cụm

Hệ gen người chứa những vùng kích thước lớn trong đó chứa các đoạn trình tự lặp lại liên tiếp nhau tạo thành một chuỗi dài một cách ngẫu nhiên. Vùng đó được gọi là ADN vệ tinh và tồn tại thành ba loại phụ thuộc vào độ dài của các cụm chứa các đoạn trình tự lặp lại (*Hình 2.9*). ADN vệ tinh là kiểu thứ nhất được xác định ở hệ gen người và gồm các cụm chứa các đoạn trình tự lặp lại có độ dài khoảng 100 đến 5000kbp. ADN tiểu vệ tinh gồm các cụm chứa các đoạn trình tự lặp lại ngắn hơn, có độ dài khoảng 100bp đến 20kbp. Ở loại thứ ba, ADN vi vệ tinh, các cụm nằm phủ lên các ADN tiểu vệ tinh, tuy nhiên, chúng được phân biệt bởi trình tự lặp lại ngắn, ít khi quá 4bp. Những trình tự lặp lại gồm hai nucleotid như [CA] là rất phổ biến và kiểu này chiếm khoảng 0,5% toàn bộ hệ gen. Những trình tự lặp lại một nucleotid chiếm khoảng 0,3%.

Các đoạn trình tự lặp lại có số lượng biến động (VNTR)

Các đoạn trình tự lặp lại liên tiếp có số lượng biến động VNTR (*v*ariable *n*umber *t*andem *r*epeats) như đoạn lặp $[CA]_n$ thường có số lượng đoạn lặp (n) biến động tại những locut cụ thể. Sự biến động này diễn ra giữa các cá thể và giữa các cặp nhiễm sắc thể tương đồng của một cá thể. Vì thế mà những trình tự đó được gọi là những đoạn lặp liên tiếp có số lượng biến động VNTR. Những đoạn trình tự lặp lại có thể nhân lên được nhờ kỹ thuật phản ứng chuỗi trùng hợp (PCR) sử dụng các đoạn mồi đặc hiệu cho những trình tự đặc thù nằm ở hai vùng biên để tạo ra nhiều sản phẩm PCR có độ dài khác nhau phụ thuộc vào số lần lặp lại của đoạn trình tự lặp (*Hình 2.11*). Độ dài của đoạn ADN nhân lên bằng PCR được dùng để phân biệt các cá thể và phân biệt các cặp nhiễm sắc

thể của một cá thể. Các VNTR có nhiều ứng dụng. Chẳng hạn như, chúng được sử dụng để xác định đặc trưng cá thể trong khoa học hình sự, xác định tội phạm từ các vật liệu sinh học tìm thấy ở hiện trường; hoặc được sử dụng trong di truyền y học để xác định những cá thể mang các gen gây bệnh di truyền.



Hình 2.11. **Các VNTR. Đoạn lặp [CA] lặp lại 4 - 16 lần. Kỹ thuật PCR dùng các đoạn mồi đặc hiệu đối với những trình tự nằm ở hai vùng biên của đoạn lặp để khuếch đại và tạo ra các đoạn ADN có độ dài khác nhau.**

II.E ĐỘT BIẾN ADN

Đột biến

Trình tự ADN của gen xác định trình tự amino acid mà nó mã hoá. Điều quan trọng là trình tự ADN phải được bảo tồn vì sự thay đổi trình tự amino acid có thể làm ảnh hưởng đến hoạt động chức năng của protein, dẫn đến những hậu quả có hại đối với sinh vật. Sự thay đổi trình tự ADN có thể xảy ra do tác động của nhiều tác nhân vật lý, hoá học và sinh học, cũng như có thể do những sai sót trong quá trình sao chép ADN. Những sự biến đổi đó được gọi là đột biến. Khi đã được sinh ra thì những biến đổi ADN được duy trì nhờ quá trình sao chép ADN và được truyền cho các tế bào thế hệ con qua quá trình phân bào. Hai thuật ngữ quan trọng dùng để mô tả các sinh vật mang đột biến là kiểu gen và kiểu hình. Kiểu gen dùng để mô tả đột biến và gen bị đột biến. Kiểu hình dùng để mô tả hậu quả đối với sinh vật. Một sinh vật biểu hiện kiểu hình thông thường (phổ biến) của loài được gọi là kiểu dại. Sinh vật có kiểu hình bị biến đổi do đột biến được gọi là thể đột biến. Đột biến xảy ra ở hai dạng: (i) đột biến điểm do sự thay đổi một base ở bất kỳ vị trí nào của gen, và (ii) đột biến thô là sự biến đổi những đoạn trình tự ADN có kích thước dài hơn. Vị trí của đột biến trong gen rất quan trọng. Chỉ những đột biến trong vùng mã hoá là ảnh hưởng đến protein và sản phẩm của gen. Đột biến ở những vùng không mã hoá hoặc ở những vùng liên gen thường không gây hậu quả.

Đột biến điểm

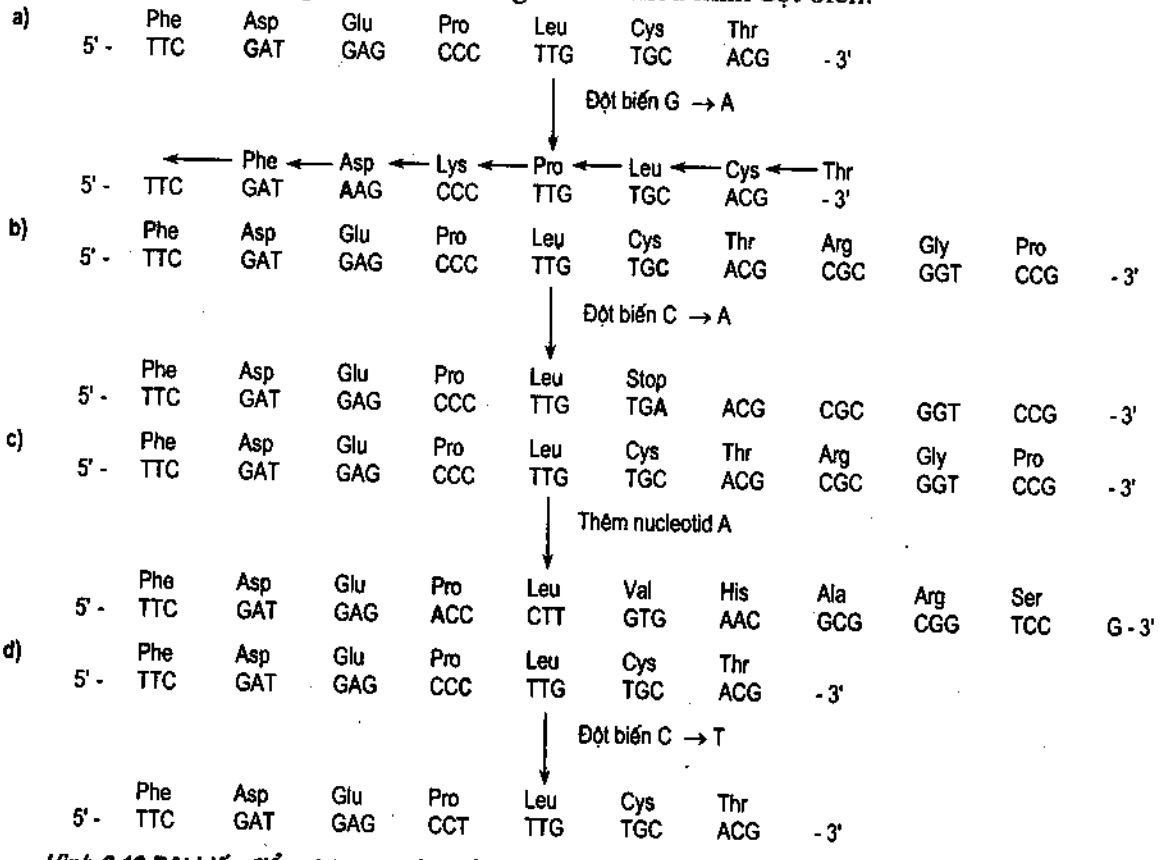
Có nhiều loại đột biến điểm, mỗi loại gây những hậu quả khác nhau đối với protein mà gen mã hoá.

Đột biến nhầm nghĩa:

Loại đột biến điểm này bao gồm những thay đổi một base dẫn đến sự thay đổi một bộ ba mã hóa cho amino acid tương ứng (*Hình 2.12a*). Những đột biến như vậy thường xảy ra ở một trong số hai base đầu của một bộ ba. Tính dư thừa của mã di truyền có nghĩa là sự thay đổi base thứ ba dường như ít gây biến đổi đến amino acid. Hậu quả của đột biến nhầm nghĩa đối với sinh vật là rất khác nhau. Hầu hết các protein không bị ảnh hưởng đối với một số biến đổi trong trình tự amino acid của nó. Tuy nhiên, những thay đổi ở những vị trí quan trọng đối với cấu trúc và chức năng của protein thường gây những hậu quả nghiêm trọng và phát sinh đột biến.

Đột biến vô nghĩa:

Đó là những đột biến điểm làm thay đổi bộ ba quy định một amino acid thành bộ ba kết thúc chuỗi (*Hình 2.12b*). Loại đột biến này làm kết thúc sớm quá trình dịch mã mARN, tạo ra protein bị cắt ngắn ở vùng chứa đầu carboxyl. Đột biến vô nghĩa thường gây hậu quả nghiêm trọng đối với hoạt tính của protein và thường sinh ra kiểu hình đột biến.



Hình 2.12. **Đột biến điểm. (a) Đột biến nhầm nghĩa, (b) Đột biến vô nghĩa, (c) Đột biến dịch khung, (d) Đột biến câm**

Đột biến dịch khung:

Các đột biến này là do sự xen thêm hoặc mất đi của các base trong gen. Nếu số lượng các base xen vào hoặc mất đi không là bội số của ba thì khung đọc sẽ bị thay đổi và ribosom sẽ đọc một bộ các bộ ba khác kể từ điểm đột biến, trong thực tế sẽ làm thay đổi trình tự amino

acid của protein được mã hoá (*Hình 2.12c*). Các đột biến dịch khung thường gây hậu quả nghiêm trọng đối với protein và sinh ra các kiểu hình đột biến.

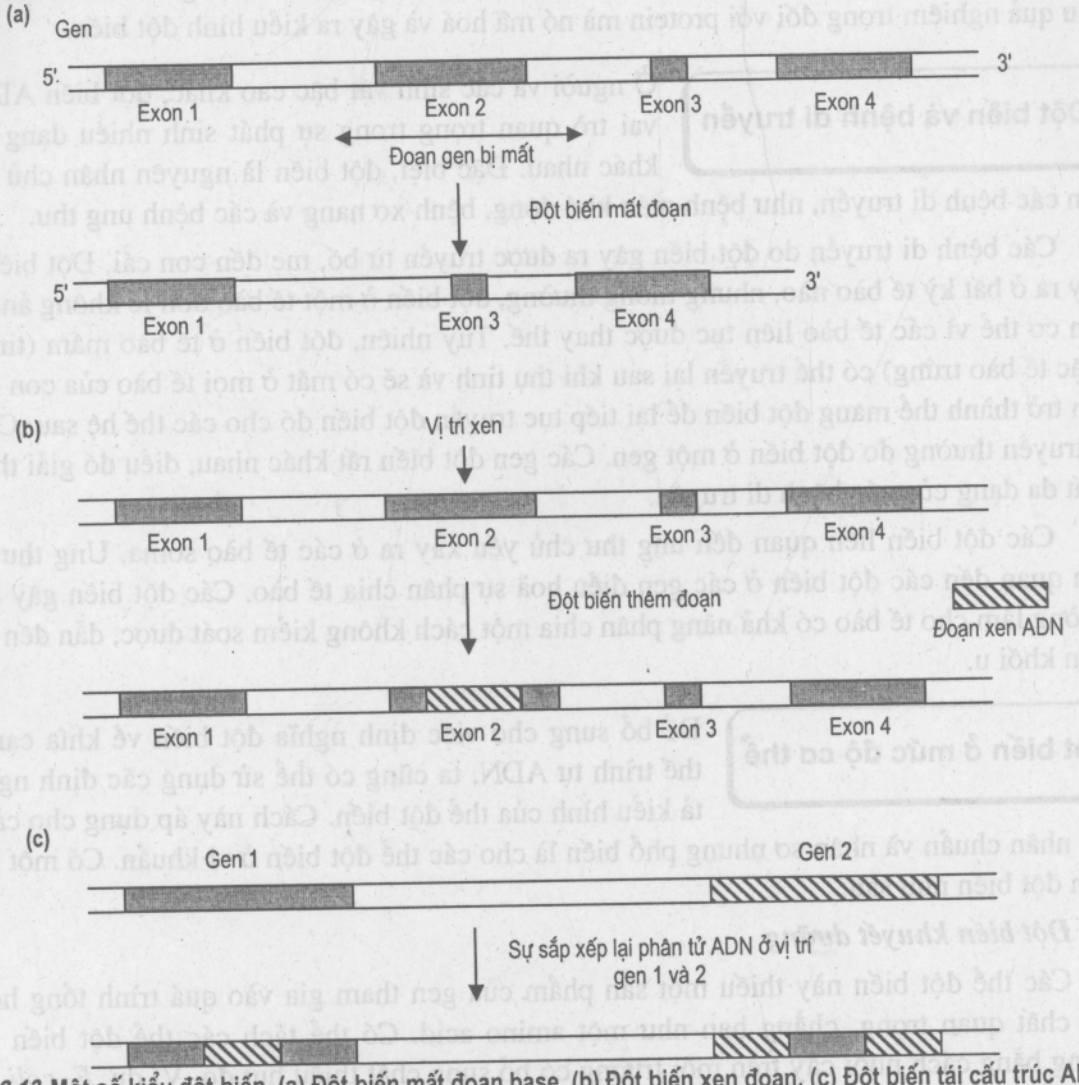
Đột biến câm:

Đột biến có thể xảy ra ở base thứ ba của một bộ ba. Nhưng, do tính dư thừa của mã di truyền mà amino acid sẽ không bị thay đổi (*Hình 2.12d*). Ngoài ra, ngay sự thay đổi của một số amino acid ở những vị trí nhất định trong chuỗi polypeptid đôi khi không làm thay đổi chức năng và hoạt tính của protein. Do vậy, đột biến câm không gây ảnh hưởng đến chức năng sinh học của protein và không sinh ra kiểu hình đột biến. Chúng tích tụ lại trong ADN của sinh vật và được gọi là các dạng đa hình. Chúng đóng góp vào tính đa dạng về trình tự ADN của những cá thể khác nhau của một loài.

Những đột biến điểm do sự thay thế base purin bằng pyrimidin hoặc ngược lại được gọi là **đột biến dị hoán.** Đột biến do thay thế nhau của hai base purin hoặc hai base pyrimidin được gọi là **đột biến đồng hoán.**

Đột biến thô

Đột biến thô gây ra những thay đổi đáng kể đối với ADN, thường ảnh hưởng đến những đoạn trình tự dài. Có nhiều loại đột biến thô.



Hình 2.13. **Một số kiểu đột biến. (a) Đột biến mất đoạn base, (b) Đột biến xen đoạn, (c) Đột biến tái cấu trúc ADN**

Mất đoạn base:

Loại đột biến này gồm sự mất một phần trình tự ADN (*Hình 2.13a*). Lượng base bị mất rất khác nhau. Đoạn base bị mất có thể nhỏ, chỉ là một base, hoặc lớn hơn nhiều. Trong một số trường hợp, đoạn trình tự mất đi có thể tương ứng với trình tự của toàn bộ một hoặc nhiều gen.

Xen đoạn base:

Trong trường hợp này, đột biến xảy ra là do xen thêm vào các base thừa, thường được bắt nguồn từ một phần khác của nhiễm sắc thể (*Hình 2.13b*). Cũng như mất đoạn base, lượng base xen vào có thể là một, hai hoặc nhiều base.

Tái cấu trúc ADN:

Loại đột biến này là sự trao đổi vị trí các đoạn trình tự ADN bên trong hoặc bên ngoài một gen với nhau (*Hình 2.13c*). Ví dụ đơn giản của loại đột biến này là đột biến xen đoạn base trong đó một đoạn trình tự ADN bị cắt ra rồi sau đó được tái kết hợp lại ở cùng vị trí, nhưng theo hướng ngược lại.

Đột biến thô, do có sự thay đổi lớn về trình tự của gen, nên dường như chắc chắn gây ra hậu quả nghiêm trọng đối với protein mà nó mã hoá và gây ra kiểu hình đột biến.

Đột biến và bệnh di truyền

Ở người và các sinh vật bậc cao khác, đột biến ADN đóng vai trò quan trọng trong sự phát sinh nhiều dạng bệnh lý khác nhau. Đặc biệt, đột biến là nguyên nhân chủ yếu gây nên các bệnh di truyền, như bệnh máu khó đông, bệnh xơ nang và các bệnh ung thư.

Các bệnh di truyền do đột biến gây ra được truyền từ bố, mẹ đến con cái. Đột biến có thể xảy ra ở bất kỳ tế bào nào, nhưng thông thường, đột biến ở một tế bào đơn lẻ không ảnh hưởng đến cơ thể vì các tế bào liên tục được thay thế. Tuy nhiên, đột biến ở tế bào mầm (tinh trùng hoặc tế bào trứng) có thể truyền lại sau khi thụ tinh và sẽ có mặt ở mọi tế bào của con làm cho con trở thành thể mang đột biến để lại tiếp tục truyền đột biến đó cho các thế hệ sau. Các bệnh di truyền thường do đột biến ở một gen. Các gen đột biến rất khác nhau, điều đó giải thích bản chất đa dạng của các bệnh di truyền.

Các đột biến liên quan đến ung thư chủ yếu xảy ra ở các tế bào soma. Ung thư thường liên quan đến các đột biến ở các gen điều hoà sự phân chia tế bào. Các đột biến gây ung thư thường làm cho tế bào có khả năng phân chia một cách không kiểm soát được, dẫn đến sự phát triển khối u.

Đột biến ở mức độ cơ thể

Để bổ sung cho việc định nghĩa đột biến về khía cạnh thay thế trình tự ADN, ta cũng có thể sử dụng các định nghĩa mô tả kiểu hình của thể đột biến. Cách này áp dụng cho cả các tế bào nhân chuẩn và nhân sơ nhưng phổ biến là cho các thể đột biến ở vi khuẩn. Có một số kiểu hình đột biến như sau:

Đột biến khuyết dưỡng

Các thể đột biến này thiếu một sản phẩm của gen tham gia vào quá trình tổng hợp một hợp chất quan trọng, chẳng hạn như một amino acid. Có thể tách các thể đột biến khuyết dưỡng bằng cách nuôi cấy trên môi trường có bổ sung chất thiếu hụt đó. Ví dụ, *E. coli* khuyết

dưỡng tryptophan không thể tổng hợp được tryptophan cho mình và phải lấy nó từ môi trường nuôi cấy. Đối ngược với các thể khuyết dưỡng là các thể nguyên dưỡng, không cần sự bổ sung dinh dưỡng đặc biệt nào.

Đột biến mẫn cảm nhiệt độ

Các thể đột biến này bị hạn chế bởi điều kiện nuôi cấy và chỉ sống được khi được nuôi cấy ở nhiệt độ nhất định. Nếu nhiệt độ tăng thì tế bào không sinh trưởng được và chết. Đột biến này thường ảnh hưởng đến một amino acid quan trọng có vai trò duy trì cấu trúc của protein. Khi nhiệt độ tăng, protein đột biến bị biến tính và mất hoạt tính, sinh ra kiểu hình đột biến.

Đột biến kháng chất kháng sinh

Các chất kháng sinh tiêu diệt các vi khuẩn kiểu dại nhưng không ảnh hưởng gì đến các thể đột biến này. Hiện tượng kháng chất kháng sinh có thể xuất hiện theo nhiều cách nhưng thường là do đột biến ở gen mã hoá cho protein đích của chất kháng sinh.

Đột biến điều hoà

Trong trường hợp này, các thể đột biến mất khả năng điều hoà sự biểu hiện của gen hay của operon. Ví dụ, ở *E. coli* tồn tại các thể đột biến ở các trình tự điều hoà sự biểu hiện các gen thuộc operon *lac* mã hóa cho các enzym cần cho sự trao đổi chất của đường lactose, làm cho những thể đột biến này có khả năng sinh tổng hợp mạnh các enzym thủy phân lactose ngay cả khi trong môi trường không có loại đường này. Các thể đột biến như vậy được gọi là các thể đột biến cơ định và thường xuất hiện do đột biến ở gen mã hóa protein ức chế *lac* (*Lac I*).

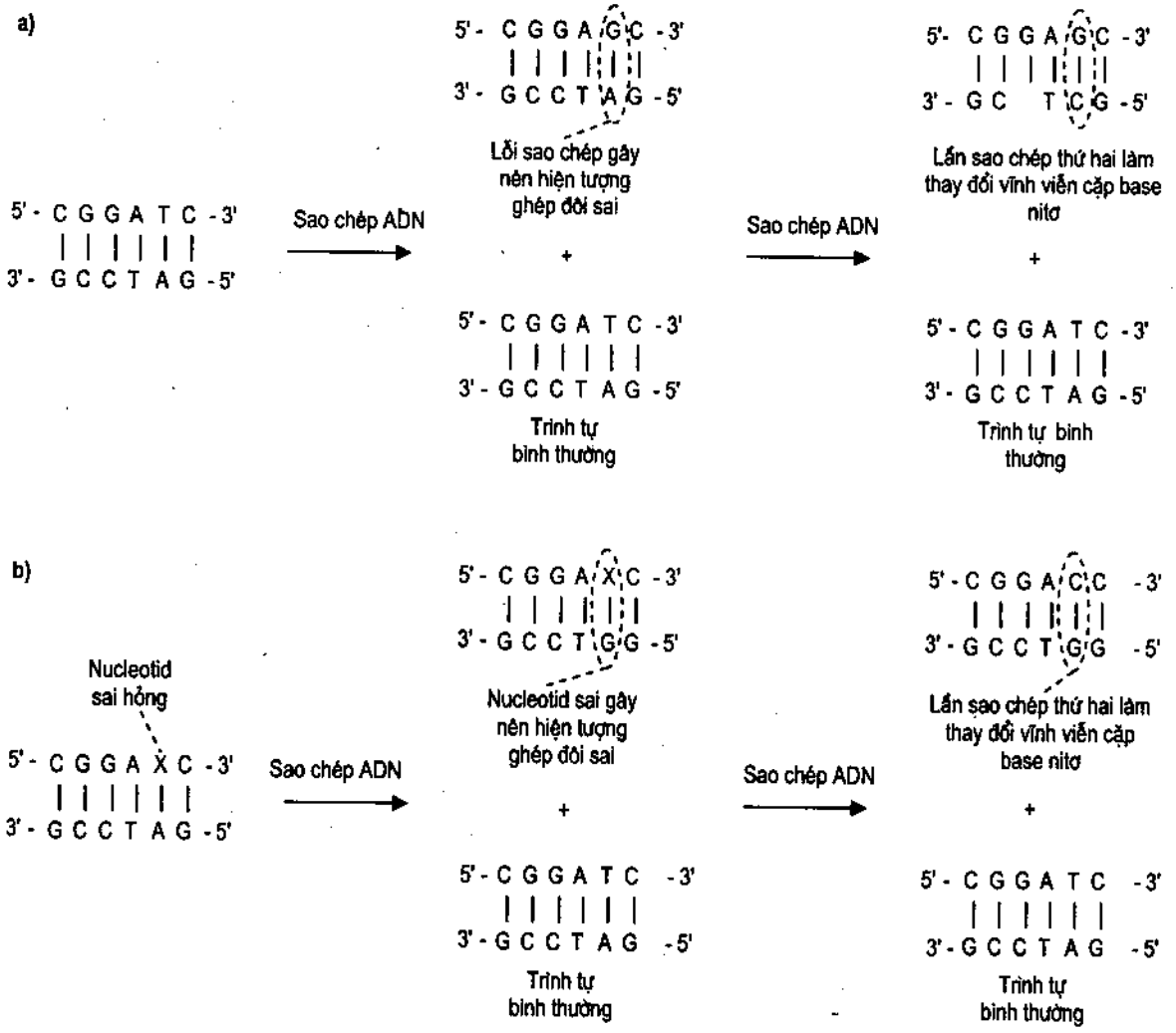
II.F CÁC TÁC NHÂN ĐỘT BIẾN VÀ SỬA CHỮA ADN

Các tác nhân gây đột biến

Đột biến có thể xuất hiện do những sai sót trong quá trình sao chép ADN. Những sai sót như vậy thường hiếm, ở vi khuẩn là khoảng $1/10^{10}$ base và có thể cao hơn ở các sinh vật bậc cao. Tỷ lệ đó được gọi là **tốc độ đột biến ngẫu nhiên** (hay đột biến tự phát). Tốc độ đột biến tăng khi tế bào chịu tác động của các **tác nhân gây đột biến** có bản chất hoá học, vật lý hay sinh học tác động trực tiếp đến ADN và làm thay đổi cấu trúc của những nucleotid đơn lẻ. Hiện tượng đó có thể dẫn đến sự thay thế hoặc phá vỡ sự kết đôi base. Chẳng hạn như khi ADN sao chép, có một base sai được xen vào đối diện với base bình thường của mạch khuôn sẽ làm thay đổi trình tự ADN. Những chu kỳ sao chép sau đó sẽ cố định sự biến đổi đó làm cho ADN bị đột biến (*Hình 2.14*). Một số tác nhân đột biến hoạt động theo cách khác như làm biến dạng nghiêm trọng phân tử ADN, gây ngăn cản sự sao chép và phiên mã.

Các hóa chất gây đột biến

Rất nhiều hoá chất, tự nhiên cũng như tổng hợp, vô cơ cũng như hữu cơ có thể tương tác với ADN, làm thay đổi cấu trúc của nó và gây ra đột biến. Hầu hết các hoá chất gây đột biến là các tác nhân gây ung thư. Ngoài ra, nhiều tác nhân vật lý cũng có thể gây đột biến, như bức xạ ion hoá dạng tia X và tia γ, bức xạ không ion hoá, đặc biệt là ánh sáng tử ngoại và cả nhiệt độ.



Hình 2.14. **Các cơ chế gây đột biến. (a) Các đột biến xuất hiện do lỗi xảy ra trong quá trình sao chép ADN; (b) Các đột biến xảy ra do sự thay đổi nucleotid dẫn đến sự thay đổi cặp base nitơ.**

Nhiều hoá chất có thể gây đột biến ADN. Một số hoá chất gây đột biến là các **đồng đẳng của base**. Các hoá chất này có cấu trúc giống với các base bình thường trong ADN và có thể lồng ghép vào ADN nhờ enzym ADN polymease khi ADN sao chép. Chúng có thể gây đột biến do làm thay đổi sự kết đôi base. Một ví dụ là **5-bromouracil** (5-BU), một đồng đẳng của thymin. 5-BU thường kết đôi với adenin; tuy nhiên, nó có thể thay đổi chút ít về cấu trúc thành dạng hỗ biến để có thể kết cặp với guanin. Sau khi ADN sao chép, cặp TA ban đầu sẽ bị thay thành cặp GC ở một trong các sợi ADN con, dẫn đến đột biến điểm (*Hình 2.15*).

Nhiều hoá chất gây đột biến khác là các **tác nhân xen**. Đó là những phân tử mỏng, phá huỷ sự sao chép ADN bằng cách xen vào giữa cặp base liền kề của chuỗi xoắn kép. Một ví dụ là **ethydium bromide**, một phân tử có chứa bốn vòng nên có cấu trúc không gian giống với một cặp base purin - pyrimidin. Ethydium bromide xen được vào chuỗi xoắn kép làm cho hai base liền kề dịch xa nhau một chút (*Hình 2.16*). Còn chưa biết chính xác bằng cách nào nó phá huỷ sự sao chép ADN, nhưng kết quả cuối cùng là nó làm cho một nucleotid xen vào vị trí của nó, gây nên đột biến dịch khung.

Nhiều tác nhân gây đột biến tác động bằng cách biến đổi các base về mặt hoá học. Trong số đó, một số tác nhân bổ sung nhóm alkyl hoặc nhóm aryl, ví dụ như **methylmethan sulfonat và ethylnitrosourea** bổ sung nhóm methyl vào nhiều vị trí của các base (*Hình 2.17*). Còn có cách biến đổi khác như sự khử amin. Ví dụ acid nitơ khử amin ở cytosin để sinh ra uracin (*Hình 2.18*) kết cặp với adenin, tạo ra sự thay thế cặp base từ GC thành AT sau khi sao chép. **Sự khử amin** của adenin thành đồng đẳng của guanin có tên là hypoxantin làm cho cặp AT bị thay thế thành cặp GC.

Đột biến có thể xảy ra ngẫu nhiên do phản ứng của ADN với một số loại hoá chất bình thường trong tế bào. Ví dụ adenin có thể ngẫu nhiên bị khử amin thành uracil. Ngoài ra, cytosin đôi khi hiển diện dưới dạng 5-methylcytosin có thể bị khử amin thành thymin. Sự biến đổi ngẫu nhiên thường xuyên xảy ra là sự khử purin do sự phá huỷ mối liên kết giữa đường deoxyribose và base purin (*Hình 2.19*). Sự có mặt của các dạng oxy phản ứng trong các tế bào ưa khí cũng có thể gây tổn thương các base. Các dạng đó bao gồm các superoxide, peroxide hydro và các gốc hydroxyl.

Các tác nhân vật lý gây đột biến

Bức xạ ion hoá cao năng như tia X và tia γ gây tổn thương các phân tử ADN, làm đứt gẫy các sợi ADN, phá huỷ đường và các base. Bức xạ không ion hoá dạng tia cực tím (UV) được các base hấp thụ và có thể sinh ra những biến đổi về cấu trúc. Đặc biệt, tia UV có thể sinh ra các cấu trúc được gọi là **cyclobutyl kép** (hay **dime cyclobutyl**) giữa các base pyrimidin liền kề nhau, đặc biệt là giữa hai base thymin (*Hình 2.20*). Sự dime hoá làm cho các base xích lại gần nhau và có thể sinh ra đột biến mất nucleotid sau khi ADN sao chép. Nhiệt cũng là một tác nhân gây đột biến quan trọng của môi trường. Hiệu ứng của nó đối với phân tử ADN là sinh ra những điểm mất base purin trên chuỗi polynucleotid, dẫn đến đột biến điểm hoặc đột biến mất nucleotid khi ADN sao chép.

Sửa chữa ADN

Nhu cầu bảo tồn trình tự ADN của gen để chống lại các tác nhân gây đột biến bắt buộc sinh vật phải phát triển các phương pháp ngăn cản đột biến bằng cách sửa chữa những thương tổn của ADN. Các cơ chế sửa chữa rất phức tạp, nhưng có thể nêu ba cơ chế chủ yếu sau đây:

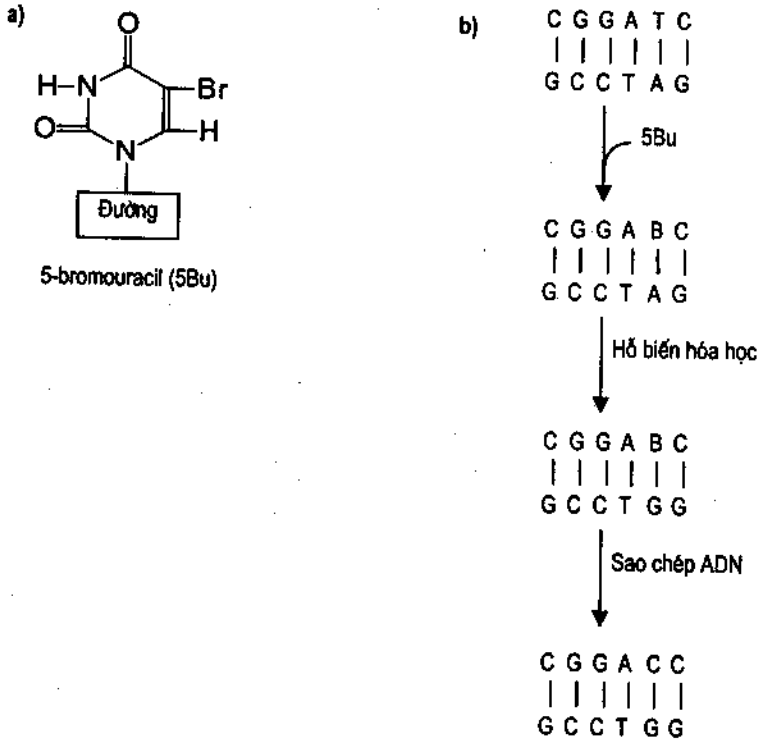
Sửa chữa bằng cắt bỏ

Đây là một hệ thống phức tạp, có lẽ là dạng sửa chữa ADN phổ biến nhất. Nhiều kiểu thương tổn ADN được sửa chữa bằng cách này, kể cả các dime pyrimidin. Đầu tiên, một trong số các enzym nhận biết các nucleotid bị thương tổn và đánh dấu nó để sửa chữa. Dấu hiệu có thể có dạng một vết khía trên một sợi của chuỗi xoắn kép nằm kề vùng bị thương tổn hoặc loại bỏ một base bị tổn thương để lại một khe hở. Ở giai đoạn sau, enzym nuclease loại bỏ nucleotid được đánh dấu và một số nucleotid xung quanh nó. Sau đó, một loại ADN polymerase (ở *E. coli* là ADN polymerase I) tổng hợp các base mới để thay thế các base bị mất, rồi ADN ligase nối đoạn ADN mới tổng hợp với sợi cũ, khôi phục lại cấu trúc ban đầu của nó (*Hình 2.21*).

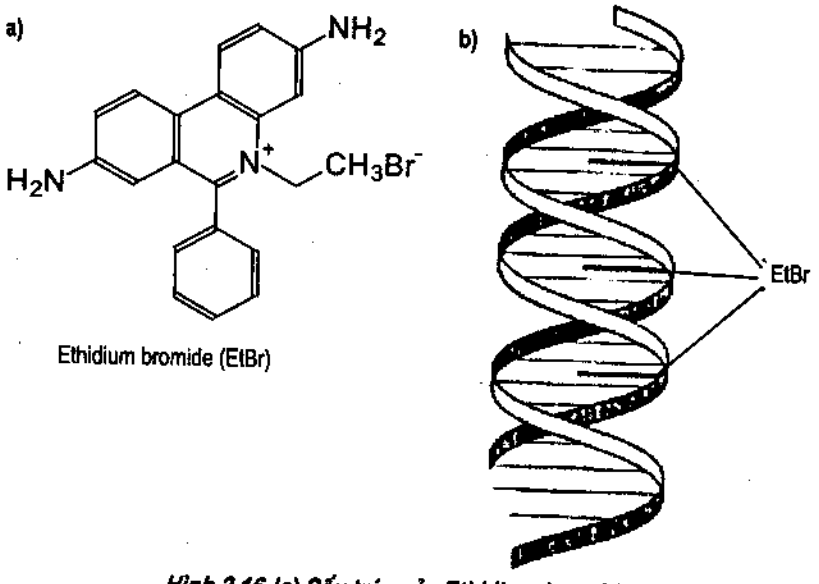
Sửa chữa trực tiếp

Dạng sửa chữa này ít phổ biến hơn, bao gồm sự biến đổi thuận nghịch về cấu trúc của nucleotid. Một ví dụ quan trọng của dạng sửa chữa này là quang phục hoạt để sửa chữa các

dime pyrimidin do bức xạ UV sinh ra. Ánh sáng nhìn thấy sẽ thúc đẩy các enzym được gọi là ADN photolyase và các enzym đó có khả năng sửa chữa các dime pyrimidin bằng cách phá vỡ các liên kết tạo nên cấu trúc dime. ADN photolyase có ở vi khuẩn, vi sinh vật nhân chuẩn và ở thực vật.



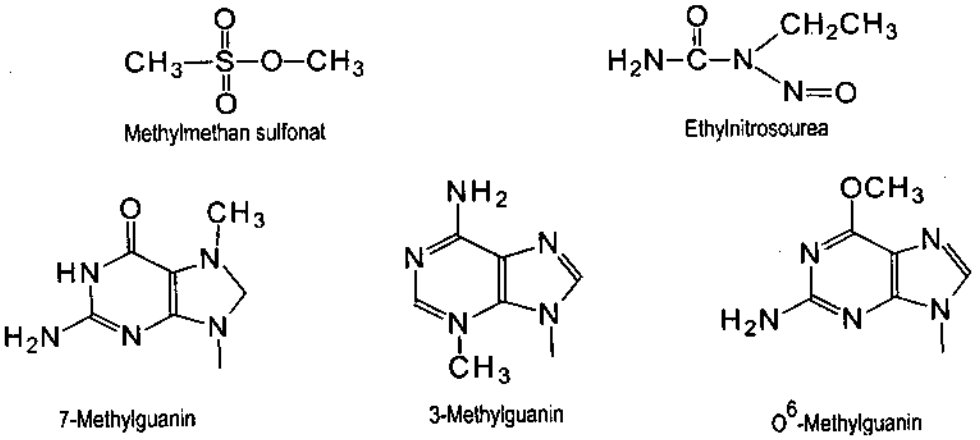
Hình 2.15. **(a) Cấu trúc của 5-bromouracil; (b) Đột biến ADN gây ra bởi 5- bromouracil**



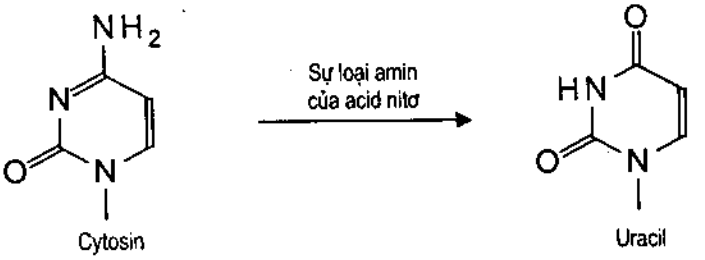
Hình 2.16. **(a) Cấu trúc của Ethidium bromide; (b) Hiệu ứng gây đột biến của Ethidium bromide**

Sửa chữa ghép đôi sai

Hệ thống này sửa chữa các sai sót trong quá trình sao chép ADN bằng cách xác định các nucleotid bị ghép đôi sai. Một số enzym đánh dấu được các chỗ ghép đôi sai, hoặc có thể trực tiếp sửa chữa chúng. Điều quan trọng là hệ thống này có thể xác định được base nào trong số base bị ghép đôi sai là base đúng và nó thực hiện điều đó bằng cách phân biệt giữa mạch ADN khuôn có trình tự đúng với sợi mới tổng hợp mang trình tự đột biến. Ở *E. coli*, mạch làm khuôn có thể nhận biết được nhờ có hiện tượng methyl hóa base adenine.

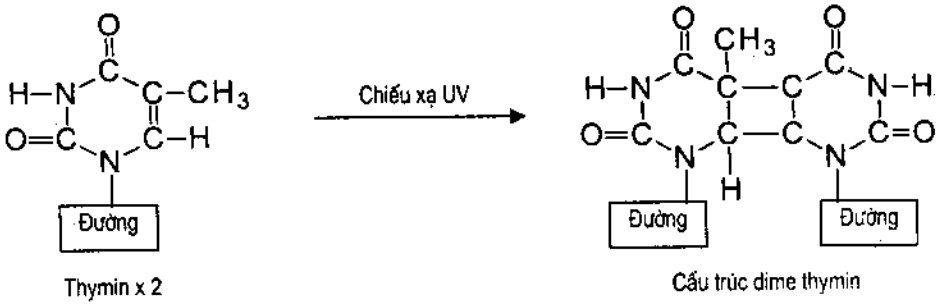


Hình 2.17. **Các hợp chất gây alkyl hóa (a) và các base nitơ sau khi bị alkyl hóa (b)**

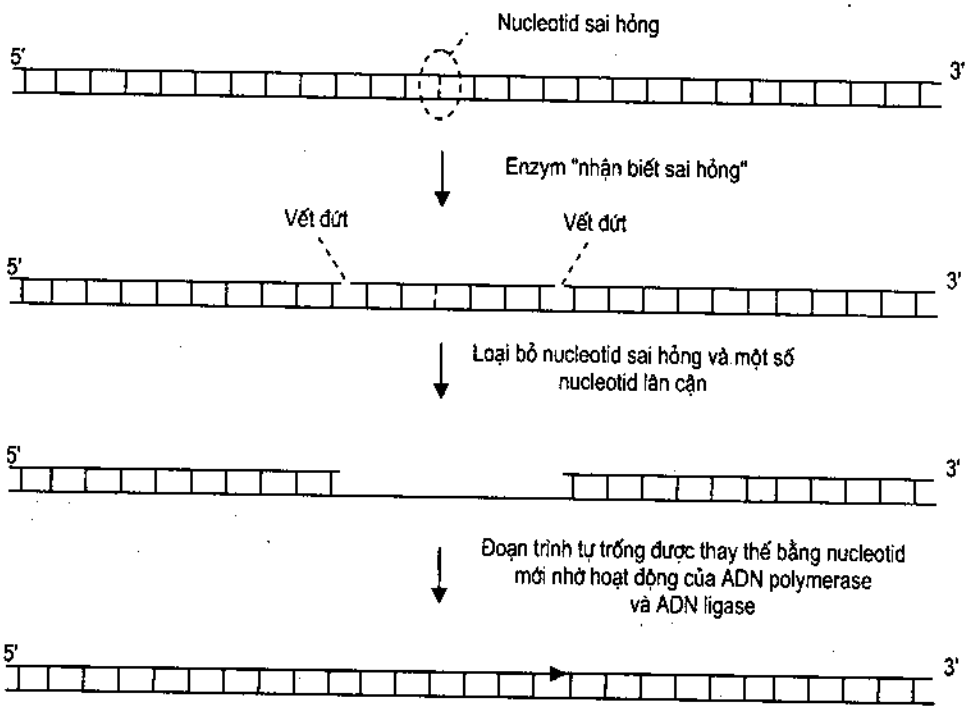


Hình 2.18. **Sự loại nhóm amin và chuyển hóa Cytosin thành Uracil gây ra bởi acid nitơ**

Hình 2.19. **Hiện tượng loại purin trong phân tử ADN**



Hình 2.20. **Sự hình thành cấu trúc thymin kép (dime thymin) do bức xạ UV**



Hình 2.21. **Sửa chữa ADN bằng cắt bỏ**

II.G TÁI TỔ HỢP

Tổng quan

Tái tổ hợp trong mục này đề cập đến sự tạo ra những tổ hợp thông tin di truyền mới trên một phân tử ADN (trên một nhiễm sắc thể). Hiện tượng này khác với sự sắp xếp lại các nhiễm sắc thể hoàn chỉnh (tức là các phân tử ADN riêng biệt) vào các giao tử và hợp tử trong quá trình giảm phân và thụ tinh. Cơ chế tái tổ hợp bao gồm sự tương tác giữa hai chuỗi xoắn kép (tức là hai phân tử ADN sợi kép). Kết quả là cả hai phân tử ADN sợi kép đứt gãy và các đầu đứt khác nhau nối lại với nhau. Cơ chế này tương đối giống với sự kiện cắt bỏ intron. Sự kiện tái tổ hợp cũng được mô tả như một trao đổi chéo, trông giống như một vắt chéo ở giảm phân. Trên quan điểm di truyền học, một nhiễm sắc tử tương đương với một chuỗi xoắn kép. Trao đổi chéo trong giảm phân xảy ra giữa hai chuỗi xoắn kép (hai nhiễm sắc tử), mỗi chiếc từ một nhiễm sắc thể trong cặp tương đồng.

Tái tổ hợp chung

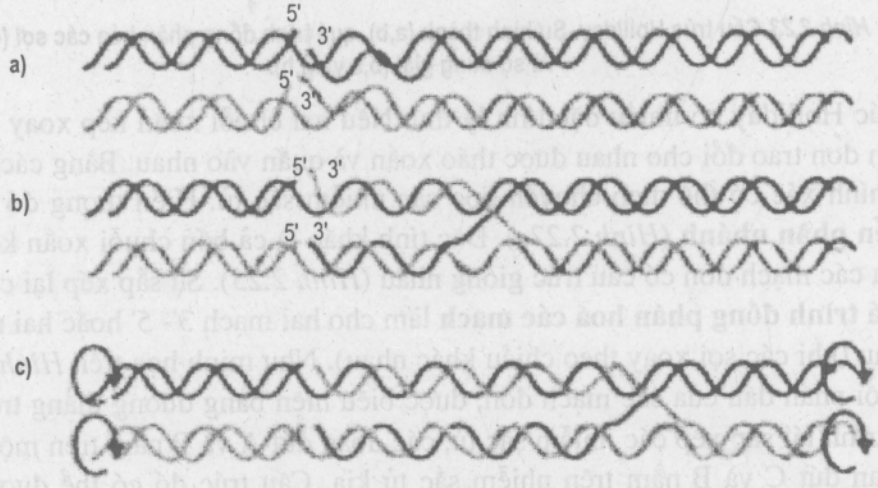
Tái tổ hợp chung là cơ chế mà các sinh vật nhân chuẩn dùng để tái tổ hợp các nhiễm sắc tử trong giảm phân. Vi khuẩn cũng dùng cơ chế đó để lồng ghép các vùng tương đồng của ADN ngoại lai vào nhiễm sắc thể của chúng sau khi ADN xâm nhập vào tế bào bằng quá trình tiếp hợp, tải nạp hoặc biến nạp. Đây là quá trình trao đổi thuận nghịch giữa hai nhiễm sắc thể. Ở tế bào nhân chuẩn, quá trình đó thường có tính bảo thủ (tức là không gì được thu nhận thêm vào hoặc mất đi, ngoại trừ sự bảo tồn). Hiệu quả là hai phân tử ADN sợi kép tương đồng nằm song song

nhau, đứt gãy ở những điểm tương ứng và mỗi đầu đứt của phân tử này được nối với phân tử kia, trao đổi toàn bộ phần còn lại của phân tử ADN sợi kép (tái tổ hợp). Như vậy, thông tin của toàn bộ trình tự được bảo tồn. Ở vi khuẩn, ADN không tạo mạch vòng ổn định sẽ bị mất. Tái tổ hợp chung cần hai phân tử ADN sợi kép có các vùng trình tự tương đồng lớn (gần giống nhau) thường dài hàng trăm cặp base. Độ dài tối thiểu của vùng tương đồng khác nhau ở các loài khác nhau.

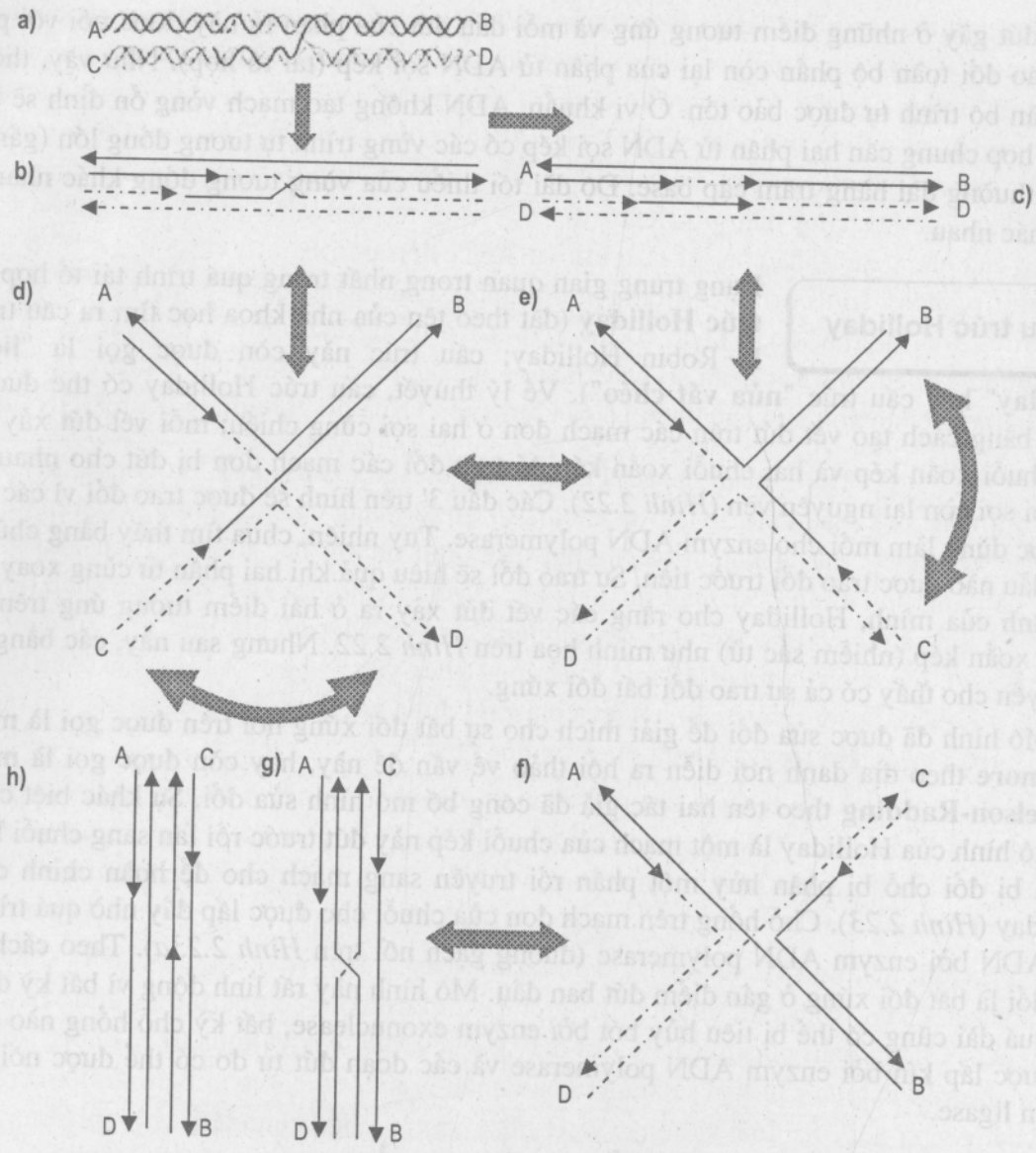
Cấu trúc Holliday

Dạng trung gian quan trọng nhất trong quá trình tái tổ hợp là **cấu trúc Holliday** (đặt theo tên của nhà khoa học tìm ra cấu trúc này là Robin Holliday; cấu trúc này còn được gọi là "**liên kết Holliday**" hay cấu trúc "**nửa vắt chéo**"). Về lý thuyết, cấu trúc Holliday có thể được hình thành bằng cách tạo vết đứt trên các mạch đơn ở hai sợi cùng chiều; mỗi vết đứt xảy ra trên một chuỗi xoắn kép và hai chuỗi xoắn kép đó trao đổi các mạch đơn bị đứt cho nhau để kết đôi với sợi còn lại nguyên vẹn (*Hình 2.22*). Các đầu 3' trên hình sẽ được trao đổi vì các đầu đó sẽ được dùng làm mồi cho enzym ADN polymerase. Tuy nhiên, chưa tìm thấy bằng chứng cho thấy đầu nào được trao đổi trước tiên. Sự trao đổi sẽ hiệu quả khi hai phân tử cùng xoay. Trong mô hình của mình, Holliday cho rằng các vết đứt xảy ra ở hai điểm tương ứng trên cả hai chuỗi xoắn kép (nhiễm sắc tử) như minh họa trên *Hình 2.22*. Nhưng sau này, các bằng chứng di truyền cho thấy có cả sự trao đổi bất đối xứng.

Mô hình đã được sửa đổi để giải thích cho sự bất đối xứng nói trên được gọi là **mô hình Aviemore** theo địa danh nơi diễn ra hội thảo về vấn đề này, hay còn được gọi là **mô hình Messelson-Radding** theo tên hai tác giả đã công bố mô hình sửa đổi. Sự khác biệt chính so với mô hình của Holliday là một mạch của chuỗi kép này đứt trước rồi lan sang chuỗi kép kia. Mạch bị đổi chỗ bị phân hủy một phần rồi truyền sang mạch cho để hoàn chỉnh cấu trúc Holliday (*Hình 2.23*). Chỗ hổng trên mạch đơn của chuỗi cho được lấp đầy nhờ quá trình tổng hợp ADN bởi enzym ADN polymerase (đường gạch nối trên *Hình 2.23a*). Theo cách đó, sự trao đổi là bất đối xứng ở gần điểm đứt ban đầu. Mô hình này rất linh động vì bất kỳ đoạn đứt nào quá dài cũng có thể bị tiêu hủy bớt bởi enzym exonuclease, bất kỳ chỗ hổng nào cũng có thể được lấp kín bởi enzym ADN polymerase và các đoạn đứt tự do có thể được nối lại nhờ enzym ligase.



Hình 2.22. (a) Sự hình thành cấu trúc Holliday bằng cách đứt một sợi và trao đổi các đoạn đứt cùng chiều. (b, c) Điểm trao đổi dịch chuyển nhờ sự xoay của chuỗi xoắn kép.



Hình 2.23. **Cấu trúc Holliday. Sự hình thành (a,b), quá trình đồng phân hóa các sợi (d,e,f) và sự dung giải (b,c và g,h).**

Cấu trúc Holliday có nhiều đặc tính lý thú. Nếu hai chuỗi xoắn kép xoay cùng hướng thì cả hai mạch đơn trao đổi cho nhau được tháo xoắn và quấn vào nhau. Bằng cách đó, điểm trao đổi chéo chính xác có thể dịch chuyển dọc các nhiễm sắc tử. Hiện tượng đó được gọi là **sự dịch chuyển phân nhánh** (*Hình 2.22c*). Đặc tính khác là cả bốn chuỗi xoắn kép tính từ điểm trao đổi của các mạch đơn có cấu trúc giống nhau (*Hình 2.23*). Sự sắp xếp lại các mạch có thể trải qua **quá trình đồng phân hoá các mạch** làm cho hai mạch 3'- 5' hoặc hai mạch 5'- 3' trao đổi cho nhau (khi các sợi xoay theo chiều khác nhau). Như minh họa trên *Hình 2.23a*, bắt đầu là sự trao đổi phần đầu của các mạch đơn, được biểu hiện bằng đường thẳng trên *Hình 2.23b*. Trong quá trình tái sắp xếp các nhiễm sắc tử, các đoạn đứt A và B nằm trên một nhiễm sắc tử, còn các đoạn đứt C và B nằm trên nhiễm sắc tử kia. Cấu trúc đó có thể được mở ra (*Hình 2.23d*), rồi xoay để các đoạn đứt C và D trao đổi với nhau, làm cho các đoạn đứt B-C và D-A không bắt chéo (*Hình 2.23e*). Lúc này cấu trúc đó mở ra cho cả bốn đoạn đứt để các đoạn đứt

B và C trao đổi cho nhau làm cho các mạch liên tục (không bị đứt, A-B và C-D) vắt chéo lên nhau (*Hình 2.23f*). Việc cắt các mạch này và nối các đoạn đứt lại sinh ra sự tái tổ hợp (trao đổi chéo, *Hình 2.23g* và *h*). Lúc này các đoạn đứt A, D trên một nhiễm sắc thể; còn C, B trên nhiễm sắc thể kia.

Như vậy, cấu trúc đó có thể được **dung giải** bằng cách cắt hai mạch ADN cùng chiều và nối lệch để tách hai chuỗi xoắn kép. Nếu các mạch đơn được trao đổi ban đầu lại bị cắt và các đoạn đứt được trao đổi cho nhau và được nối lại thì phân tử ADN sợi kép ban đầu được khôi phục, không sinh ra tái tổ hợp và những đầu đứt cũ vẫn sẽ nối với nhau (*Hình 2.23c*). Nếu hai mạch kia cũng bị đứt và trao đổi cho nhau thì cả hai mạch đơn của cả hai chuỗi xoắn kép sẽ bị cắt và trao đổi cho nhau. Đó là hiện tượng đứt gãy sợi kép chéo nhau và gây ra tái tổ hợp các gen ở cả hai phía, do các đoạn đứt của các phân tử ADN sẽ trao đổi cho nhau. Đó là lý do vì sao hiện tượng này được gọi là tái tổ hợp.

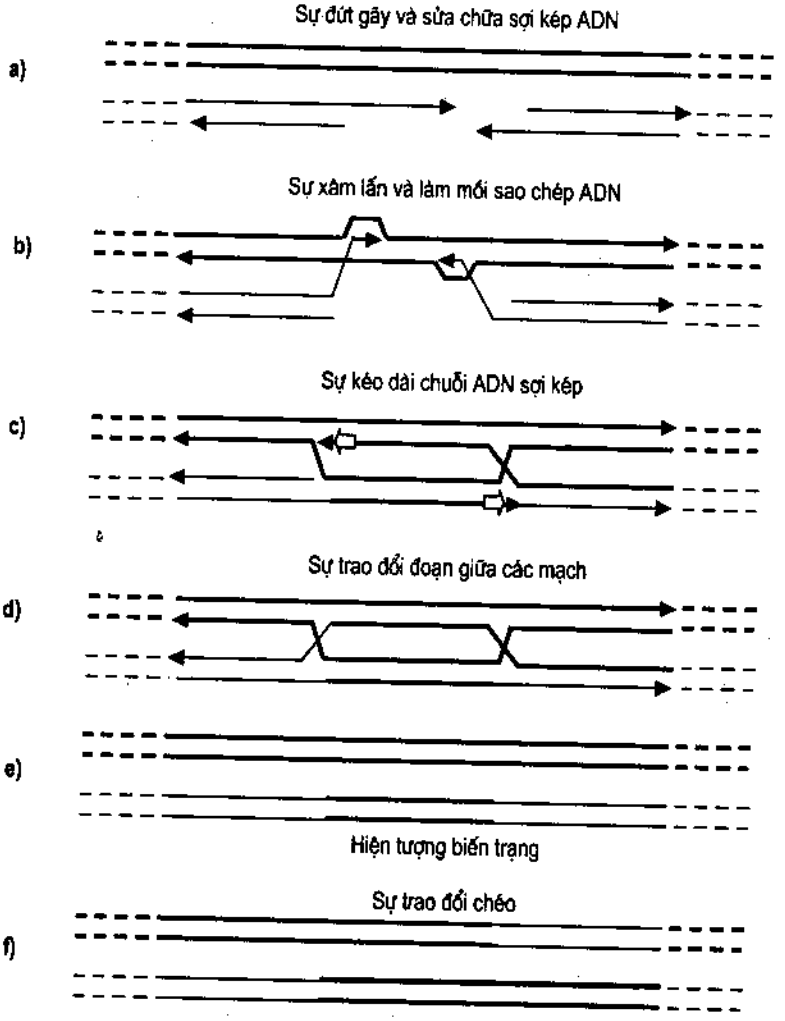
Mô hình đứt gãy sợi kép

Mô hình đứt gãy sợi kép để tái tổ hợp được xây dựng để giải thích những đặc điểm tái tổ hợp quan sát được ở nấm men *S. cereviseae*. Rất có thể là mô hình này hoạt động ở mọi sinh vật nhân chuẩn và mô hình Aviemore thực sự không diễn ra trong tự nhiên. Tuy nhiên, vẫn còn có ít bằng chứng liên quan đến cơ chế thực sự hoạt động ở các sinh vật nhân chuẩn có hệ gen lớn. Một số nhà khoa học tin rằng tính phức tạp và kích thước hệ gen của hầu hết các sinh vật nhân chuẩn đòi hỏi một cơ chế tái tổ hợp thận trọng hơn so với cơ chế đứt gãy sợi kép. Sự tái tổ hợp ở *S. cereviseae* được khởi đầu bằng việc sinh ra vết đứt ở một chuỗi xoắn kép (*Hình 2.24a*). Protein SPO11 thực hiện chức năng này có liên quan tới loại enzym topoisomerase II và hình thành nên liên kết đồng hoá trị ADN-protein qua vết đứt. Đầu 5' bị phân hủy ngược dòng với điểm gãy, để lại đoạn gãy dài khoảng 500bp từ đầu 3', có thể phát hiện được bằng phương pháp sinh hóa. Mô hình giả thiết rằng các đoạn đứt đầu 3' xâm lấn chuỗi xoắn kép ADN tương đồng còn nguyên vẹn (*Hình 2.24b*) và có thể làm mồi để tổng hợp qua chỗ đứt nhờ enzym ADN polymerase, trong đó dùng mạch bổ trợ của ADN sợi kép còn nguyên vẹn làm khuôn (*Hình 2.24c*). Khi chỗ đứt đã được nối, sự trao đổi các mạch tiếp sau sẽ sinh ra hai cấu trúc Holliday, mỗi cấu trúc ở một bên của chỗ đứt ban đầu (*Hình 2.24d*). Nếu cả hai cấu trúc đó là đồng tính (nghĩa là các mạch đơn được cắt rồi nối lại như cũ) thì sẽ không có hiện tượng tái tổ hợp (*Hình 2.24e*), nhưng trình tự tại vết đứt ở nhiễm sắc tử bị đứt sẽ biến trạng giống với trình tự ở nhiễm sắc thể kia. Nếu hai cấu trúc Holliday không đồng tính thì sẽ có sự trao đổi chéo ở một phía nào đó của vết đứt và các đoạn đứt của hai chuỗi xoắn kép sẽ tái tổ hợp (*Hình 2.24f*).

Sự biến trạng

Sự biến trạng là hiện tượng sinh ra tỷ lệ phân ly không theo quy luật Mendel ở bốn tế bào đơn bội được tạo ra bởi một tế bào phân chia giảm nhiễm một lần. Nếu toàn bộ thông tin di truyền được bảo tồn thì hai trong số bốn tế bào đơn bội sẽ mang bản sao của một alen, còn hai tế bào kia sẽ mang bản sao của alen còn lại, tương ứng với quy luật Mendel. Đôi khi nguyên tắc này bị phá vỡ và ta có **tỷ lệ phân ly 3:1** (thay cho 2:2). Ở một số nấm, do có thêm một lần nguyên phân nên các sản phẩm đơn bội của giảm phân sẽ được nhân đôi, biến tỷ lệ 3:1 thành 6:2. Cũng có khi sinh ra **tỷ lệ 5:3**, là biến dạng của **tỷ lệ 4:4** hoặc sinh ra tỷ lệ hiếm gặp là **7:1**. Trong mỗi trường hợp trên, các tế bào chị em sinh ra từ nguyên phân khác nhau về mặt di truyền; sự kiện đó được

gọi là **sự phân ly sau giảm phân hay PMS** (*post meiotic segregation*; được mô tả ở phần dưới đây). Các chuỗi xoắn kép ADN được cắt nối trong quá trình tái tổ hợp phải chứa những cặp base bị ghép đôi sai (*Hình 2.25*, chú ý rằng ở vùng trao đổi bất đối xứng chỉ có ở một trong số bốn nhiễm sắc tử có thể mang đôi base ghép sai.)

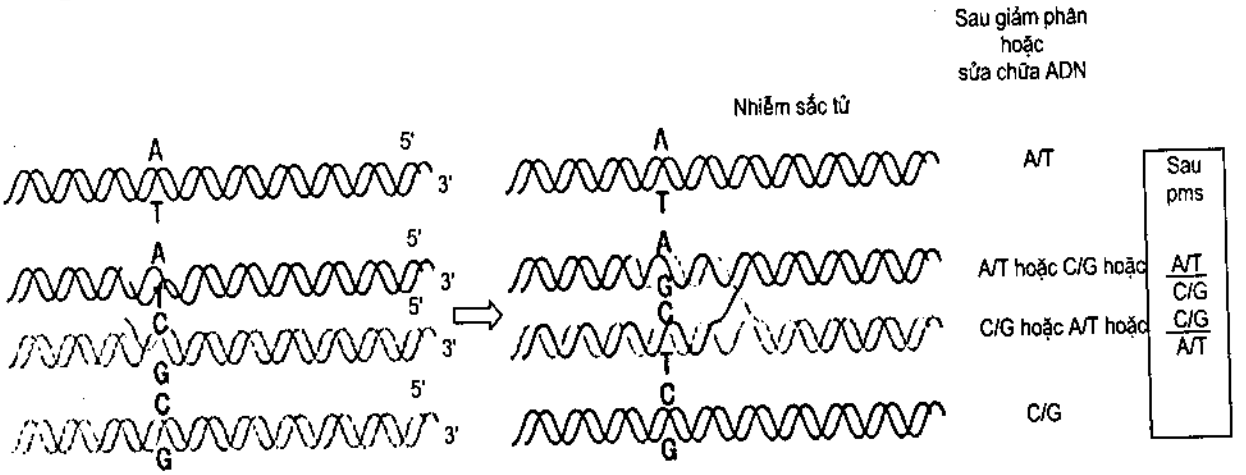


Hình 2.24. **Mô hình tái tổ hợp do đứt gãy sợi kép. (a) Vết đứt gãy sợi kép, (b) Các đoạn đứt tự do xâm lấn sang phân tử ADN tương đồng còn nguyên vẹn. Chúng mở rộng qua khe gãy (c, hướng mũi tên) và một cấu trúc Holliday hình thành ở mỗi phía của vết đứt (d). Cơ chế trên có thể giải thích cho việc hình thành sự biến trạng (e) hoặc tái tổ hợp (f).**

Giả sử một nhiễm sắc thể trong cặp nhiễm sắc thể bố, mẹ có cặp AT trong khi cặp nhiễm sắc thể kia có cặp GC. Nếu các mạch đơn ADN trao đổi nhau ở vùng này thì **chuỗi kép dị hợp** (do mỗi mạch đơn ADN bắt nguồn từ một nhiễm sắc thể) sẽ có một cặp AG và cặp CT. Nếu các **cặp ghép đôi sai** này được enzym sửa chữa nhận ra, chúng có thể được sửa đổi thành AT hoặc thành GC, theo một trong hai cách đó (hoặc hoàn toàn không được sửa chữa) làm cho ba trong số bốn nhiễm sắc tử có thể mang trình tự như nhau. Nếu các đôi base ghép sai không được sửa chữa trước lần tổng hợp ADN tiếp theo thì hai mạch sẽ được dùng làm khuôn để sao chép thành các chuỗi kép mới. Các chuỗi kép đó sẽ tách nhau trong nguyên phân sinh ra các tế bào con mang các trình tự khác nhau, một mang AT và một mang GC. Vì điều đó xảy ra ở

nguyên phân và sau quá trình giảm phân, nên nó được gọi là **sự phân ly sau giảm phân**. Có ba khả năng xảy ra đối với mỗi nhiễm sắc tử, hai khả năng do hướng sửa chữa, và một là không được sửa chữa, vì thế có chín tổ hợp có khả năng được tạo ra, một trong các khả năng đó khôi phục lại sự sắp xếp như ở bố, mẹ. Chuỗi kép dị hợp có thể mang đôi ghép sai được trình bày trên *Hình 2.23* và *Hình 2.24*. Sự trao đổi đối xứng mà không có sự sửa chữa các đôi ghép sai sẽ sinh ra tỷ lệ biến đổi 4:4 trong quá trình phân ly sau giảm phân.

Sự biến trạng sinh ra tỷ lệ 3:1 sau giảm phân dễ dàng được giải thích bằng mô hình đứt gãy chuỗi kép vì các base ở chỗ đứt trên nhiễm sắc tử bị đứt luôn luôn được thay thế dựa vào thông tin về trình tự từ nhiễm sắc thể tương đồng kia.



Hình 2.25. **Sự biến trạng gen do sửa đổi ghép đôi sai ở chuỗi kép dị hợp ADN**

Các enzym tái tổ hợp

Sự tái tổ hợp cần nhiều protein. Một số protein tương tự như protein RecA của *E. coli*. Các protein này gắn vào các mạch đơn ADN và xen vào các trình tự ADN sợi kép tương đồng. Các protein ở vi khuẩn, nấm men và người rất giống nhau. Phần lớn các enzym tham gia vào quá trình sửa chữa ADN cũng hoạt động chức năng trong quá trình tái tổ hợp ở giảm phân. Sự sửa chữa vết đứt gãy ở chuỗi kép do bức xạ ion hoá rất hiệu quả ở nấm men và đòi hỏi sự tái tổ hợp ở mức độ cao đúng như dự đoán của mô hình sửa chữa vết đứt ở chuỗi kép. Các enzym sửa chữa bằng cắt bỏ cũng hoạt động mạnh trong quá trình tái tổ hợp ở giảm phân. Phức hệ nhận biết các đôi base ghép sai để định vị các lỗi sai sót như các dime thymin sau khi chiếu tia UV dường như cũng hoạt động trong quá trình giảm phân. Một số protein sửa chữa ghép đôi sai cũng có mặt ở các giai đoạn ghép đôi và nối của các nhiễm sắc thể tương đồng, và có thể giúp hạn chế sự tái tổ hợp xảy ra giữa các ADN tương đồng. Việc loại bỏ các đôi base ghép sai có thể ngăn cản sự tái tổ hợp giữa các nhiễm sắc thể ở con lai giữa các loài có quan hệ họ hàng, tạo nên sự cách ly sau giao phối của các loài đó. Các protein sửa chữa khác khoanh vùng khu vực tái tổ hợp. Các gen cần cho quá trình tái tổ hợp cũng thường là các gen quan trọng cho quá trình sửa chữa ADN bình thường. Khi chúng hoạt động sai thì những sai hỏng ADN có thể gây nên đột biến mới, có thể dẫn đến ung thư. Chẳng hạn sản phẩm thông thường của gen gây ung thư vú ở người, gen *BRCA1*, là một thành phần của phức hệ tiếp hợp có vai trò trong việc kết đôi và tái tổ hợp các nhiễm sắc thể trong giảm phân. Các gen bị sai hỏng

trong trường hợp bệnh giãn mao mạch mất điều hòa là những gen quan trọng tham gia vào quá trình tái tổ hợp trong giảm phân.

Tái tổ hợp tại điểm đặc hiệu

Phage lambda (λ) có thể lồng ghép ADN của nó vào nhiễm sắc thể vật chủ *E. coli* ở trạng thái tiềm tan (trạng thái nghỉ) và sẽ sao chép cùng ADN vật chủ (trong giai đoạn gây tan). Việc lồng ghép được thực hiện nhờ trao đổi chéo xảy ra giữa 15 cặp base được gọi là điểm gắn *att* trên nhiễm sắc thể vật chủ và trình tự tương tự trên nhiễm sắc thể mạch vòng của phage (điểm *att* còn được gọi là *'O'*). Hiện tượng này xảy ra cần hoạt động của enzym intergrase (được gọi tắt là *Int*) do hệ gen phage mã hoá. Việc cắt bỏ cũng xảy ra theo kiểu tái tổ hợp và do enzym excisionase xúc tác (gen *xis* của phage mã hoá enzym này).

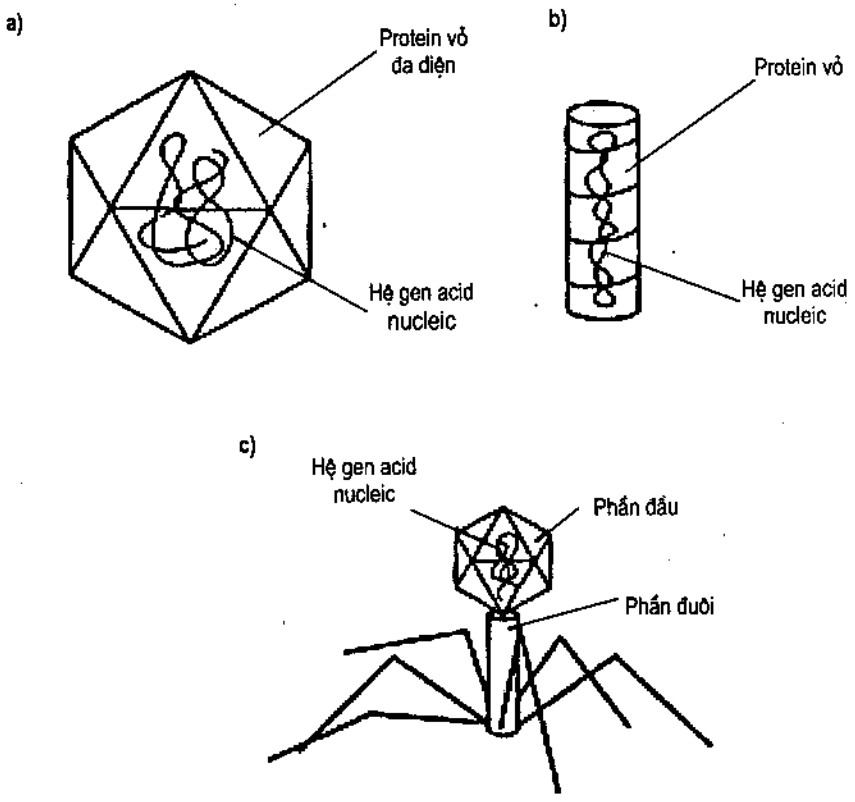
II.H PHAGE

Phage

Các phage (phagơ) là những virus lây nhiễm vi khuẩn. Nhiều loại phage có xu hướng lây nhiễm đặc hiệu, nghĩa là chỉ lây nhiễm những vi khuẩn chủ nhất định. Cũng giống như các virus khác, chúng được cấu tạo gồm hệ gen là acid nucleic ở bên trong và được bao bọc bằng lớp vỏ protein bảo vệ ở bên ngoài. Có ba dạng phage như minh họa trên *Hình 2.26*: (i) dạng đa diện, có các phân tử polypeptid tạo thành một cấu trúc hình học bao bọc lấy acid nucleic. Ví dụ: phage M2 gây nhiễm *E. coli*; (ii) dạng nhiều sợi nhỏ hoặc xoắn, trong đó các chuỗi polypeptid được sắp xếp như một chuỗi xoắn tạo thành cấu trúc hình que. Ví dụ: phage M13 gây nhiễm *E. coli*; (iii) dạng có đầu và đuôi, là dạng cấu trúc phức tạp nhất, gồm đầu có cấu trúc đa diện và đuôi dạng sợi để acid nucleic của phage xâm nhập vào tế bào chủ. Ví dụ về dạng này là phage T4 và λ lây nhiễm ở *E. coli*. Acid nucleic của phage có thể là ADN hoặc ARN, mạch đơn hay sợi kép. Thông thường toàn bộ hệ gen đều nằm trên một phân tử acid nucleic duy nhất, tuy nhiên ở một số phage ARN, các gen nằm trên nhiều hơn một phân tử acid nucleic và các phage như vậy được gọi là các phage phân đoạn. Hệ gen phage có biến động lớn về kích thước, từ vài kbp đến khoảng 150kbp. Số lượng gen thay đổi tương ứng với kích thước hệ gen. Hệ gen phage M13 dài khoảng 6kbp và chỉ chứa 10 gen. Các phage có hệ gen dài hơn, đặc biệt là các phage có cấu trúc phức tạp (như có đầu và đuôi) có nhiều gen hơn. Ví dụ, phage T4 dài 166kbp, chứa 150 gen. Các gen phage mã hoá cho các protein cần thiết cho việc xây dựng lớp vỏ capsid và các enzym tham gia vào quá trình sao chép ADN của phage. Tất cả các phage đều cần một số protein và các loại ARN của vật chủ. Một đặc tính của một số phage (như ΦX174) là các gen nằm gối lên nhau (được dịch mã theo những khung đọc khác nhau). Hiện tượng này cũng bắt gặp ở các virus của tế bào nhân chuẩn.

Chu trình sống của phage

Sau khi phage lây nhiễm tế bào chủ, những sự kiện xảy ra trong tế bào chủ có thể khác nhau, song chủ yếu theo hai hướng là gây tan hoặc tiềm tan. Ở con đường gây tan, phage nhanh chóng làm tan tế bào chủ sau khi lây nhiễm. Ở con đường tiềm tan, phage chỉ làm tan tế bào chủ sau một giai đoạn "yên lặng". Những phage khác nhau có các cách lây nhiễm khác nhau. Các phage nhiễm theo kiểu gây tan được gọi là các phage độc. Các phage nhiễm theo kiểu tiềm tan được gọi là các phage ôn hòa.



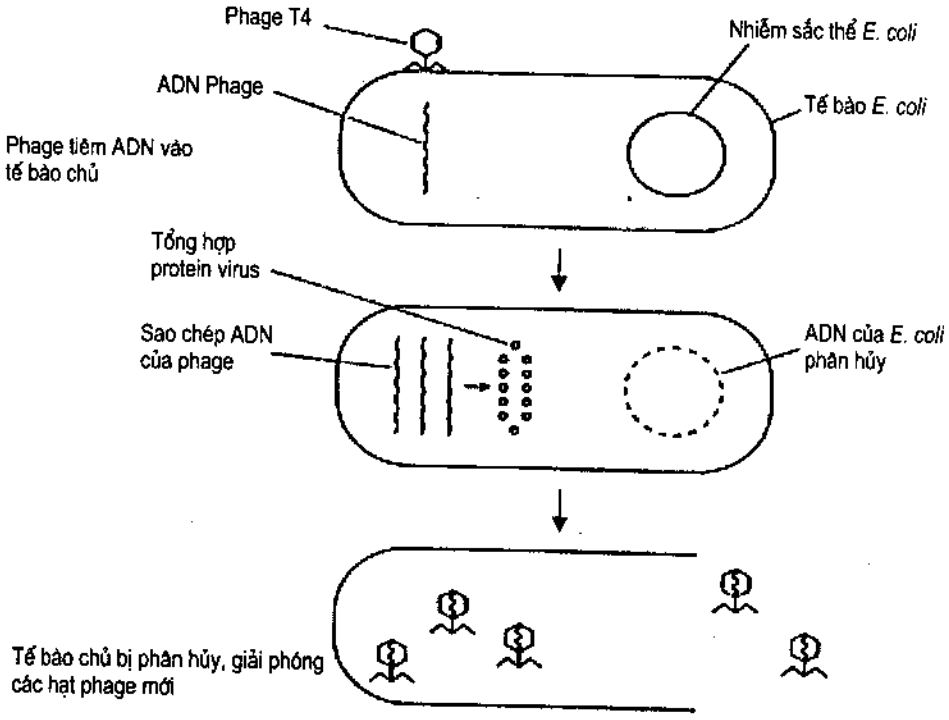
Hình 2.26. **Cấu trúc của phage. (a) dạng đa diện, (b) dạng sợi hoặc xoắn, (c) dạng có cấu trúc đầu và đuôi**

Con đường gây tan của phage T4

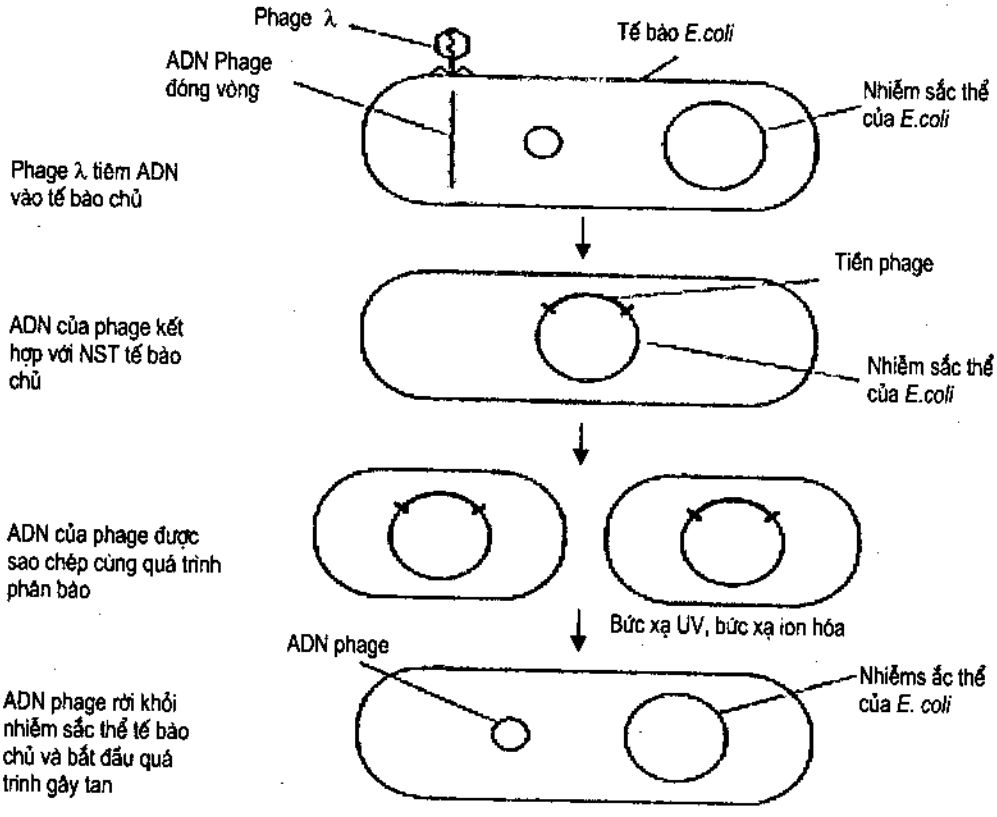
Phage T4 lây nhiễm gây tan *E. coli* (*Hình 2.27*). Các hạt phage gắn với protein thụ thể trên bề mặt tế bào (được gọi là protein *omp*C) và protein phage được tiêm vào tế bào qua phần đuôi. Trong tế bào, các quá trình phiên mã các gen của phage và sao chép ADN bắt đầu. Lúc này, ARN và các protein của tế bào chủ bị khống chế. ADN của tế bào chủ bị khử mạch polypeptid và các nucleotid của nó được dùng để sao chép ADN của phage. Các protein vỏ capsid của phage được tổng hợp và các phage mới được lắp ráp. Cuối cùng, tế bào chủ bị tan ra và khoảng 200-300 hạt phage mới được giải phóng ra.

Chu trình tiềm tan của phage lambda (λ)

Phage ôn hoà λ có thể nhiễm gây tan ở *E. coli*, nhưng thường là nó nhiễm tiềm tan (*Hình 2.28*). Sau khi phage tiêm ADN của nó vào tế bào, ADN của nó hình thành cấu trúc mạch vòng (nhờ có trình tự *cos* gồm 12 nuleotid bổ trợ với nhau ở hai đầu) rồi lồng ghép vào nhiễm sắc thể vật chủ. Sự lồng ghép diễn ra theo cơ chế trao đổi chéo tại đoạn trình tự gồm 15 cặp base tương đồng với đoạn trình tự trên nhiễm sắc thể của *E. coli*. Sự lồng ghép luôn diễn ra ở một vị trí nhất định, và dạng lồng ghép như vậy được gọi là tiền phage. Tiền phage được duy trì như vậy nhiều thế hệ và được sao chép cùng với nhiễm sắc tế bào chủ. Cuối cùng, sau nhiều lần phân bào nó chuyển sang dạng nhiễm gây tan. Sự chuyển trạng thái này có thể được cảm ứng bởi nhiều tác nhân vật lý và/hoặc hoá học. Nhìn chung, các tác nhân đó đều gây tổn thương ADN và có thể là tín hiệu báo cái chết của tế bào chủ sắp đến gần. Để đáp lại các kích thích đó, ADN phage tách khỏi tế bào chủ nhờ cơ chế trao đổi chéo thuận nghịch. ADN phage sao chép và bao gói vào vỏ capsid. Cuối cùng, tế bào vi khuẩn tan ra và các hạt phage mới được giải phóng ra.



Hình 2.27. **Sự nhiễm gây tan *E. coli* bởi phage T4**



Hình 2.28. **Sự nhiễm tiềm tan *E. coli* bởi phage λ.**

Phage M13

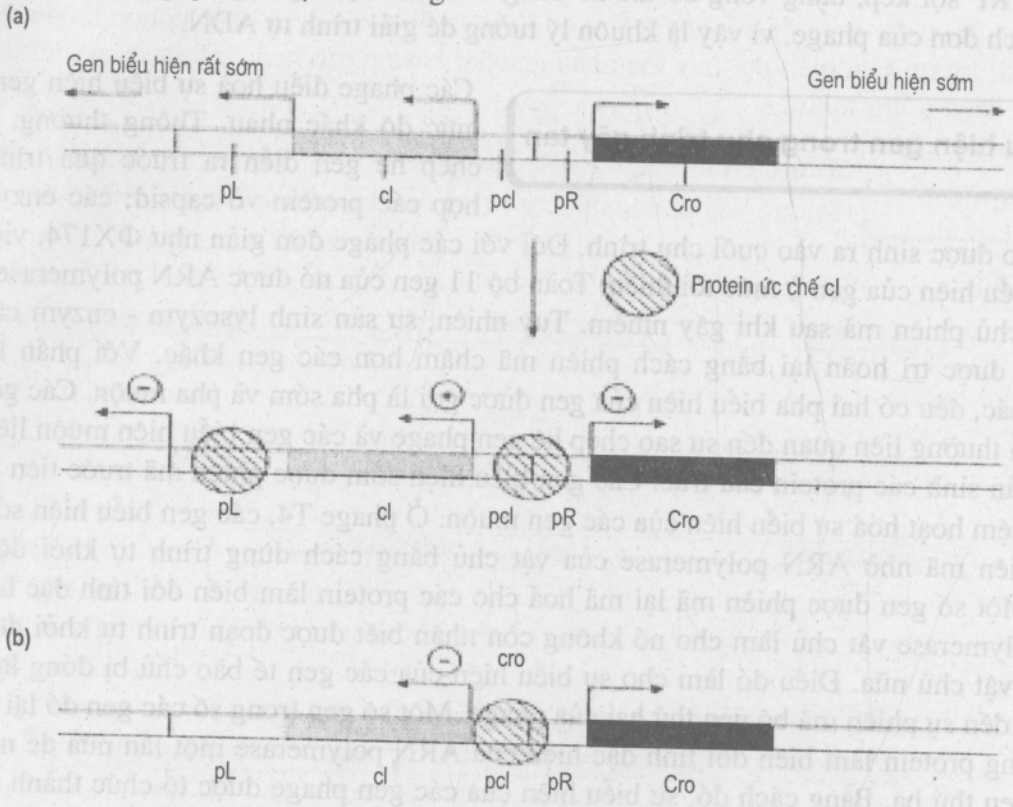
Một số phage có sự biến động giữa chu trình tiềm tan và gây tan. Ví dụ như phage M13 gây nhiễm ở *E. coli*. Hệ gen của nó là phân tử ADN mạch đơn, dạng vòng. Khi nhiễm, các enzym của tế bào vật chủ sẽ tạo ra dạng sao chép (dạng RF, *r*eplication *f*orm) có cấu trúc sợi kép. Nhiều bản sao của RF được tổng hợp làm khuôn để tạo ra các dạng mạch đơn mới sẽ được bao gói trong vỏ capsid và được giải phóng khỏi tế bào mà không làm tan tế bào. Các hạt phage cũng còn được truyền cho các tế bào khác thông qua quá trình phân bào và các tế bào bị nhiễm M13 tiếp tục sinh trưởng và phân chia sinh ra các tế bào con bị nhiễm. Nhiều đặc điểm trong chu trình sống của phage M13 làm cho nó trở thành một vector tách dòng hiệu quả trong các nghiên cứu sinh học phân tử. Dạng RF sợi kép, dạng vòng có thể dễ dàng làm tinh sạch giống như đối với plasmid và ADN mạch đơn của phage, vì vậy là khuôn lý tưởng để giải trình tự ADN.

Sự biểu hiện gen trong chu trình gây tan

Các phage điều hoà sự biểu hiện gen ở các mức độ khác nhau. Thông thường, sự sao chép hệ gen diễn ra trước quá trình tổng hợp các protein vỏ capsid; các enzym gây tan tế bào được sinh ra vào cuối chu trình. Đối với các phage đơn giản như ΦX174, việc điều hoà sự biểu hiện của gen ở mức tối thiểu. Toàn bộ 11 gen của nó được ARN polymerase của tế bào vật chủ phiên mã sau khi gây nhiễm. Tuy nhiên, sự sản sinh lysozym - enzym cần cuối cùng, sẽ được trì hoãn lại bằng cách phiên mã chậm hơn các gen khác. Với phần lớn các phage khác, đều có hai pha biểu hiện của gen được gọi là pha sớm và pha muộn. Các gen biểu hiện sớm thường liên quan đến sự sao chép hệ gen phage và các gen biểu hiện muộn liên quan đến sự sản sinh các protein cấu trúc. Các gen biểu hiện sớm được phiên mã trước tiên và chịu trách nhiệm hoạt hoá sự biểu hiện của các gen muộn. Ở phage T4, các gen biểu hiện sớm nhất được phiên mã nhờ ARN polymerase của vật chủ bằng cách dùng trình tự khởi động của phage. Một số gen được phiên mã lại mã hoá cho các protein làm biến đổi tính đặc hiệu của ARN polymerase vật chủ làm cho nó không còn nhận biết được đoạn trình tự khởi động của các gen vật chủ nữa. Điều đó làm cho sự biểu hiện của các gen tế bào chủ bị đóng lại, đồng thời dẫn đến sự phiên mã bộ gen thứ hai của phage. Một số gen trong số các gen đó lại mã hoá cho những protein làm biến đổi tính đặc hiệu của ARN polymerase một lần nữa để nó phiên mã bộ gen thứ ba. Bằng cách đó, sự biểu hiện của các gen phage được tổ chức thành các pha với những sản phẩm của các gen sớm làm nhiệm vụ mở các gen muộn.

Phage λ được nghiên cứu tích cực như một mô hình về sự biểu hiện các gen ở phage. Khi nhiễm vào *E. coli*, phage λ có thể gây tan hoặc tiềm tan. Có ba pha biểu hiện của các gen phage: rất sớm, sớm và muộn. Sự biểu hiện của các gen sớm được điều hoà bởi hai trình tự khởi động *pL* và *pR*, nằm ở hai bên gen điều hoà *cI* mã hoá protein ức chế (*Hình 2.29*). Trong quá trình gây tan, sự phiên mã các gen khởi động *pL* và *pR* P_L đều bị khống chế bởi protein của gen *cI*; protein này cũng đồng thời kích thích sự biểu hiện của chính gen mã hoá nó khi nó gắn lên P_R và đồng thời phủ lên một phần trình tự khởi động của *cI*. Khi có mặt protein cI, sự biểu hiện của các gen sớm và muộn đều bị ức chế và sẽ có sự nhiễm gây tan. Sự lựa chọn giữa con đường gây tan và tiềm tan là do protein Cro của phage λ, một loại protein hoạt động như chất ức chế sự phiên mã của gen *cI*. Sau khi nhiễm, ARN polymerase của vật chủ bắt đầu phiên mã các gen λ, gen *cI* được biểu hiện và ngăn cản sự phiên mã các gen sớm, khống chế tiến trình gây tan. Gen *Cro* cũng được biểu hiện và nếu lượng protein Cro được tích luỹ đủ, nó

sẽ loại bỏ sự khống chế của các gen sớm, cho phép sự nhiễm gây tan tiếp tục tiến triển. Sự cạnh tranh giữa các protein cI và Cro quyết định tế bào lựa chọn con đường gây tan hoặc tiềm tan dường như là một quá trình ngẫu nhiên, chỉ phụ thuộc vào protein nào được tích luỹ nhanh hơn. Cho đến khi nào lượng protein cI còn đủ thì quá trình tiềm tan vẫn được duy trì. Tuy nhiên, nếu mức protein cI giảm thì sự gây tan sẽ được cảm ứng. Sự cảm ứng gây tan sau một giai đoạn tiềm tan có thể được kích thích bởi các tác nhân như bức xạ ion hoá, đây là tác nhân gây hoạt hoá cơ chế bảo vệ chung ở *E. coli* được gọi là cơ chế sửa chữa kiểu SOS. Quá trình này cần gen *RecA* mà protein của nó có hoạt tính bất hoạt chất ức chế cI bằng cách phân cắt protein này. Khi cI bị bất hoạt, các gen biểu hiện sớm được hoạt hoá, quá trình tiềm tan kết thúc và con đường gây tan được cảm ứng.



Hình 2.29. **Kiểm soát sự biểu hiện gen ở phage λ. (a) chất ức chế cI khống chế sự phiên mã các gen sớm và kích thích sự phiên mã gen cI, (b) chất ức chế Cro khống chế sự phiên mã gen cI, cho phép các gen sớm được biểu hiện.**

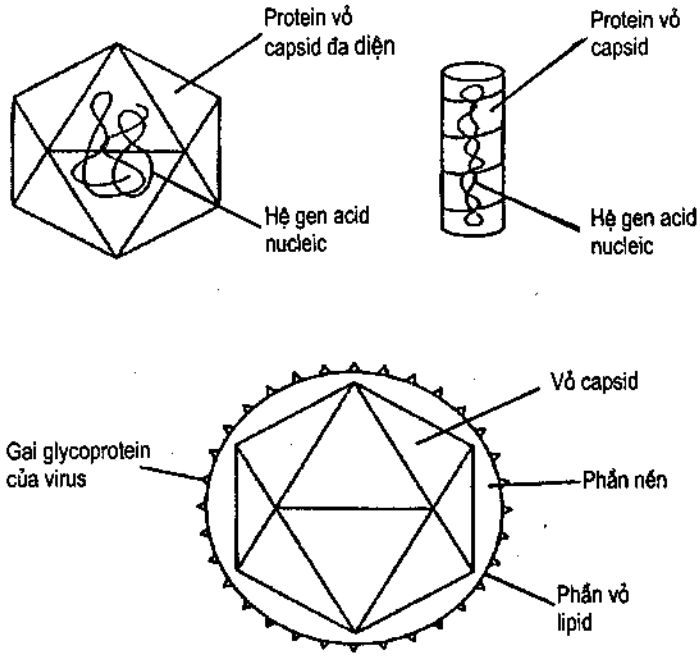
II.I VIRUS LÂY NHIỄM SINH VẬT NHÂN CHUẨN

Virus

Virus là các tác nhân gây nhiễm siêu hiển vi, cấu tạo gồm hệ gen là acid nucleic được bao bọc bởi vỏ protein. Chúng là các vật ký sinh và sinh sản bằng cách gây nhiễm tế bào và sử dụng khả năng sao chép ADN và tổng hợp protein của chính tế bào chủ. Rất khó coi chúng là những cơ thể sống vì chúng không có khả năng tồn tại độc lập. Rất nhiều virus là các tác nhân gây bệnh cho người và các sinh vật khác. Ngoài sinh vật nhân chuẩn, các sinh vật nhân sơ như vi khuẩn cũng có thể bị nhiễm virus.

Hệ gen virus

Có nhiều loại virus gây nhiễm các tế bào nhân chuẩn. Chúng khác nhau về hệ gen acid nucleic: có thể là ADN sợi đơn (ssDNA), ADN sợi kép (dsDNA), ARN sợi đơn (ssRNA) hoặc ARN sợi kép (dsRNA). Acid nucleic có thể có dạng thẳng hoặc vòng. Trong một số trường hợp, hệ gen virus chỉ là một phân tử ADN duy nhất. Còn ở những trường hợp khác, các gen virus tồn tại trên nhiều hơn một phân tử ADN và hệ gen như vậy được gọi là hệ gen phân đoạn. Các hệ gen sợi đơn có thể dương tính hoặc âm tính. Việc xác định trình tự dương tính và âm tính liên quan đến trình tự của gen và các trình tự mARN được phiên mã từ nó. Nếu hệ gen mang nghĩa dương tính thì trình tự của nó sẽ giống với trình tự mARN được phiên mã. Ngược lại, nếu hệ gen âm tính thì trình tự của nó có trình tự bổ trợ với mARN. Trong một số trường hợp, hệ gen mã hoá cho các mARN mang cả hai nghĩa. Nhiều hạt virus không có hệ gen hoàn chỉnh hoặc khả năng hoạt động chức năng nên chúng chỉ sao chép khi được hỗ trợ bởi loại virus trợ giúp kiểu dại có khả năng sao chép được lây nhiễm đồng thời vào tế bào chủ hoặc nhờ quá trình tác động tương hỗ giữa các dạng đột biến khác nhau cùng đồng thời lây nhiễm tế bào chủ.



Hình 2.30. **Cấu trúc virus lây nhiễm ở sinh vật nhân chuẩn. (a) dạng cấu trúc đa diện, (b) dạng sợi hoặc xoắn, (c) dạng có lớp vỏ lipid và gai glycoprotein**

Cấu tạo của virus

Hệ gen virus được bao bọc bởi một lớp vỏ bảo vệ được gọi là vỏ capsid hình thành từ các chuỗi polypeptid do chính hệ gen virus mã hoá. Tổ hợp gồm hệ gen và vỏ capsid được gọi là nucleocapsid. Nucleocapsid có hai hình dạng cơ bản (*Hình 2.30*): (i) dạng đa diện, trong trường hợp này, các phân tử polypeptid hình thành một cấu trúc hình học bao bọc lấy acid nucleic, ví dụ như proliovirus; (ii) dạng sợi hoặc xoắn; trong trường hợp này, các chuỗi polypeptid được sắp xếp như một chuỗi xoắn để hình thành một cấu trúc hình que bao lấy hệ gen acid nucleic. Ở nhiều loại virus, vỏ capsid lại được bao bởi vỏ lipid có nguồn gốc từ màng tế bào vật chủ khi nó được giải phóng ra khỏi tế bào. Glycoprotein do virus mã hoá và các protein có nguồn gốc từ tế bào vật chủ có thể được xen vào lớp màng này. Glycoprotein của virus đóng vai trò quan

trọng trong việc xúc tiến sự lây nhiễm bằng cách tương tác với các protein thụ thể trên bề mặt tế bào vật chủ. Các protein nền cho phép nucleocapsid và vỏ lipid tương tác với nhau. Các protein nền và protein vỏ capsid còn có chức năng khác nữa liên quan đến sự sao chép và phiên mã các gen của virus. Các protein cấu trúc của virus thường mang tính kháng nguyên và là đặc điểm cần quan tâm trong quá trình phát triển vaccin.

Các chiến lược sao chép

Virus tạo ra các bản sao của chính nó bằng cách khai thác khả năng sao chép và tổng hợp protein của tế bào vật chủ. Các virus sử dụng những chiến lược khác nhau để sao chép và phiên mã các gen của mình. Mọi virus đều dựa vào tế bào chủ để dịch mã và xác định quy mô phiên mã cũng như sao chép. Các virus ADN phụ thuộc vào các enzym tế bào vật chủ nhiều hơn các virus ARN, và các virus có hệ gen kích thước nhỏ phụ thuộc vào bộ máy phiên mã và sao chép của tế bào vật chủ nhiều hơn các virus có hệ gen kích thước lớn. Các virus ARN cần enzym ARN polymerase sử dụng ARN làm mạch khuôn vốn không có trong các tế bào vật chủ mà do chính virus mã hoá để tiến hành quá trình sao chép. Các virus ARN khác được gọi là retrovirus sao chép thông qua dạng trung gian là ADN sợi kép được sinh ra từ ARN của virus nhờ enzym phiên mã ngược (reverse transcriptase, enzym này chỉ thấy ở hệ gen virus mã hoá) có khả năng phiên mã để chuyển trình tự ARN thành ADN.

Mức độ nhạy cảm của tế bào đối với sự lây nhiễm của virus liên quan đến sự xuất hiện và khả năng "hỗ trợ" virus của các thụ thể đặc hiệu đối với từng loại virus trên bề mặt tế bào chủ. Những tế bào có khả năng hỗ trợ virus lây nhiễm và sao chép được gọi là các tế bào cho phép.

Sự lây nhiễm virus không phải lúc nào cũng gây biểu hiện bệnh. Chức năng trước hết của virus là sao chép bản thân hệ gen của nó, hậu quả của sự kiện đó là có thể đồng thời gây tổn thương cho tế bào. Nếu tế bào bị tổn thương nghiêm trọng có thể gây nên bệnh. Khả năng gây bệnh của virus được gọi là tính độc, và đó cũng là điều bất lợi đối với virus vì nó có thể làm giảm khả năng sao chép và sinh sản của chính nó.

Virus ADN

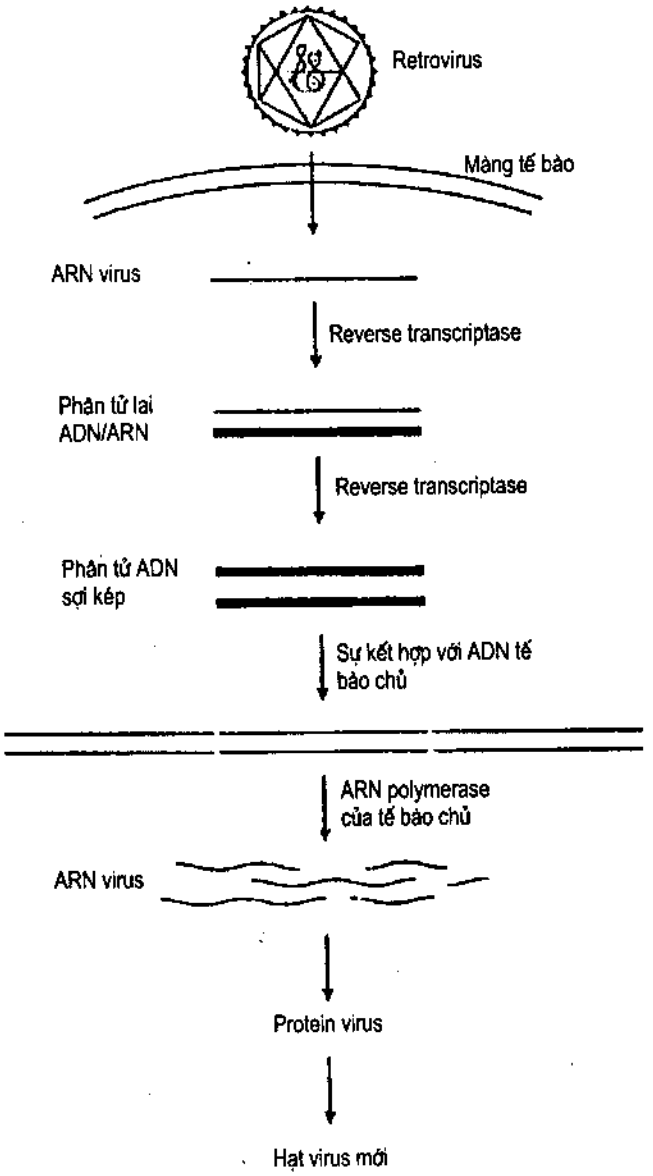
Hệ gen virus ADN có sự biến động lớn về kích thước, từ 5kbp đến 270kbp. ADN có thể có cấu trúc sợi đơn hoặc sợi kép, dạng thẳng hoặc vòng. Hầu hết các virus ADN sao chép trong nhân tế bào vật chủ, sử dụng bộ máy sao chép hệ gen của tế bào chủ và sự biểu hiện của gen virus. Các virus ADN kích thước lớn mã hoá khoảng 200 gen và có chu trình sống phức tạp với hình thức biểu hiện các gen và sự sao chép hệ gen được điều hòa. Một ví dụ của virus ADN kích thước lớn là virus bệnh mụn rộp, gây ra nhiều bệnh như thủy đậu, zona, chứng viêm các tuyến bạch hầu cho nhiều động vật có xương sống (kể cả ở người). Virus mụn rộp dạng 1 gây bệnh hecpet môi là một loại virus mụn rộp được nghiên cứu kỹ. Nó có hệ gen ADN sợi kép dài 150kbp chứa hơn 70 gen ở cả trên hai mạch của phân tử ADN, một số gen nằm gối lên nhau. Sự phiên mã và sao chép hệ gen virus được điều hòa chặt chẽ. Các gen virus chia thành ba nhóm được gọi là nhóm gen biểu hiện rất sớm (α), sớm (β) và muộn (γ). Thời gian biểu hiện của chúng diễn ra theo một trình tự xác định sau khi lây nhiễm vào tế bào chủ. Sự biểu hiện của gen còn được điều hoà cùng với sự sao chép hệ gen.

Các virus ADN kích thước nhỏ có rất ít gen, vì vậy mà sự sao chép và biểu hiện gen của chúng phụ thuộc nhiều hơn vào tế bào vật chủ. Ví dụ như virus bại liệt và virus sinh khối u ở khỉ SV40 thuộc họ virus papova. Virus SV40 có hệ gen nhỏ (5kbp) chứa 5 gen. Các gen được tổ chức trong một hệ gen nhỏ bằng cách được mã hoá trên cả hai mạch của ADN và có trình tự nằm gối lên nhau. Các protein virus được sinh ra bằng cách sử dụng các khung đọc khác nhau

và cắt nối luân phiên các đoạn trình tự. Các gen được biểu hiện thành hai pha là pha sớm và pha muộn. Các gen biểu hiện sớm sinh ra các protein hoạt hóa quá trình phiên mã, gọi là các kháng nguyên T kích thước lớn và các kháng nguyên T kích thước nhỏ. Các protein đó kích thích sự phiên mã và sao chép cả hai hệ gen tế bào chủ và virus, đồng thời quy định các đặc tính sinh khối u của virus.

Virus ARN

Hệ gen virus ARN có thể ở dạng mạch đơn hay sợi kép. Nếu là mạch đơn thì hệ gen có thể mang nghĩa dương tính hoặc âm tính. Trong một số trường hợp, hệ gen mã hoá các mARN mang cả hai nghĩa. Sự phiên mã các gen virus từ hệ gen ARN cần enzym ARN polymerase sử dụng ARN làm mạch khuôn vốn không có sẵn ở tế bào chủ mà do gen *pol* của virus mã hoá. Vì các virus ARN không cần enzym ADN polymerase của tế bào vật chủ nên sự phiên mã và sao chép không nhất thiết phải xảy ra trong nhân như đối với các virus ADN.



Hình 2.31. **Chu trình sống của retrovirus**

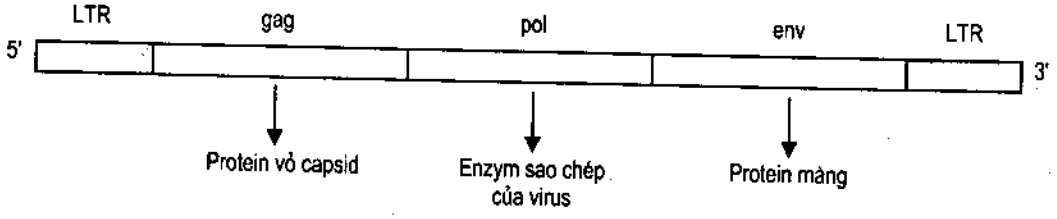
Như vậy, có nhiều virus ARN sao chép ở tế bào chất. Không như các polymerase của tế bào chủ, ARN polymerase sử dụng ARN làm khuôn không có khả năng đọc sửa và sao chép khuôn mẫu của nó với rất nhiều sai sót. Tốc độ đột biến là $1/10^3$ - 10^4/một chu kỳ sao chép, cao hơn rất nhiều so với tốc độ đột biến ở các virus ADN là khoảng $1/10^8$ - 10^{11}. Hệ quả của tốc độ đột biến cao là chúng có khả năng tiến hoá nhanh và có thể biến đổi tính kháng nguyên và tính độc, điều đó cho phép chúng nhanh chóng thích nghi với sự biến đổi của môi trường và chống lại những nỗ lực của hệ thống miễn dịch ở vật chủ nhằm loại bỏ chúng.

Một số virus ARN đột biến nhanh tới mức chúng tồn tại như một quần thể gồm nhiều trình tự hệ gen, sao chép bằng cách tác động bổ trợ lẫn nhau và được gọi là loài *quasi*. Nhiều đột biến là có hại đối với quá trình sao chép của virus ARN. Chính điều đó đã tạo ra giới hạn cho kích thước hệ gen virus chỉ vào khoảng 10^4 nucleotid, một kích thước chỉ vừa đủ để sinh ra trung bình một đột biến ở một hệ gen.

Retrovirus

Đây là một nhóm virus ARN quan trọng có hệ gen là phân tử ARN sợi đơn mang nghĩa dương tính. Nhóm này bao gồm cả virus gây suy giảm miễn dịch ở người (HIV), chúng gây nên hội chứng suy giảm miễn dịch mắc phải (AIDS). Retrovirus chứa hai bản sao ARN trong mỗi hạt virus. Khi gây nhiễm tế bào vật chủ, ARN sợi đơn xâm nhập vào tế bào và được chuyển thành ADN sợi kép nhờ hoạt động của enzym phiên mã ngược. Phân tử ADN sợi kép của virus sau đó được lồng ghép vào hệ gen tế bào vật chủ nhờ enzym intergrase của virus. Dạng lồng ghép như vậy được gọi là tiền virus (provirus) và sau đó hoạt động như khuôn mẫu để sao chép hệ gen virus và biểu hiện các gen của virus (*Hình 2.31*). Hệ gen của các retrovirus đều có cấu trúc tương tự và có một số trình tự tương đồng với các đoạn trình tự trên hệ gen người có tên là các yếu tố di truyền vận động kiểu phiên mã ngược (retrotransposon). Mức độ phức tạp của các retrovirus rất khác nhau. Một số rất đơn giản và chỉ khác retrotransposon ở chỗ có thêm gen *env* mã hoá cho glycoprotein vỏ cần cho hoạt tính lây nhiễm. Các loại khác, như HIV, có hệ gen kích thước lớn hơn với chu trình sống phức tạp hơn, mã hoá cho các protein hoạt động ở các giai đoạn khác nhau của chu kỳ sao chép và tham gia vào việc điều hòa chức năng của cả virus và tế bào chủ.

Một số virus gây ung thư ở động vật (nhưng ít liên quan đến ung thư ở người). Các retrovirus này được gọi là các tác nhân gây ung thư và thường mang gen, được gọi là gen gây ung thư mà tế bào vật chủ mắc phải do tái tổ hợp với hệ gen virus. Các gen gây ung thư mã hoá cho các protein tham gia vào quá trình điều hoà sự sinh trưởng của tế bào và có bản sao của nó được gọi là gen tiền ung thư. Sự biểu hiện của các gen gây ung thư là tế bào phân chia không kiểm soát được dẫn đến sự hình thành khối u.



Hình 2.32. **Hệ gen retrovirus**

Hệ gen của các retrovirus thường có cấu trúc cơ bản giống nhau (*Hình 2.32*). Hầu hết hệ gen retrovirus chứa ba gen là *gag*, *pol* và *env*. Gen *gag* mã hóa cho protein vỏ capsid, *pol* mã hoá cho các enzym tham gia vào quá trình sao chép hệ gen virus (gồm enzym phiên mã ngược, ARNaseH, intergrase và protease), và *env* mã hoá cho các protein có ở lớp vỏ lipid. Ở hai đầu của hệ gen có đoạn trình tự lặp lại được gọi là các đoạn trình tự lặp lại ở các đầu mút (LTR) tham gia vào quá trình sao chép, lồng ghép và biểu hiện các gen của virus. Sự phiên mã các gen virus từ dạng tiền virus phụ thuộc vào ARN polymerase II của tế bào vật chủ. Enzym này gắn với LTR ở đầu 5' để khởi đầu quá trình phiên mã. Các mARN được dịch mã thành các polyprotein để rồi được cắt thành các protein của virus bằng cách phân giải protein. Một số retrovirus như HIV sinh thêm các mARN bằng cách cắt nối chọn lọc. Enzym phiên mã ngược của một số retrovirus có mức độ nhầm lẫn cao. Điều đó có nghĩa là nhiều bản sao hệ gen của virus được sinh ra không còn khả năng tự sao chép. Hiện tượng này được khắc phục bằng cách hai bản sao hệ gen trong một hạt virus tác động bổ trợ cho nhau nhờ tái tổ hợp trong quá trình phiên mã ngược. Những đặc tính đó cùng với tốc độ đột biến và thay thế cao (10^9 - 10^{10} hạt virus mới/ngày) giúp cho HIV-1 thích nghi được với những điều kiện môi trường mới như áp lực của chọn lọc do thuốc kháng sinh hoặc điều trị.

Phần III
CÁC CƠ CHẾ DI TRUYỀN

Phần III gồm các mục:

III.A Di truyền học Mendel
- Các khái niệm cơ bản
- Phép lai đơn tính
- Phát hiện thể dị hợp tử

Những tỷ lệ biến đổi của tỷ lệ 3:1
Đa alen

III.B Di truyền học Mendel mở rộng
- Phép lai hai tính
- Tương tác át chế

Các quy luật Mendel
Kiểm tra số liệu thực nghiệm

III.C Giảm phân và quá trình phát sinh giao tử
- Sự sinh sản
- Giảm phân
- Giảm phân và tái tổ hợp

Sự phát sinh giao tử
Sự phát sinh các giao tử dị bội

III.D Liên kết gen
- Định nghĩa
- Cách đo lường đơn giản sự liên kết gen
- Phép lai ba tính

Nhiễu
Phân tích liên kết gen ở nấm
Tần số tái tổ hợp và khoảng cách vật lý

III.E Sự truyền gen ở vi khuẩn
- Tổng quan
- Plasmid là trung gian truyền gen

Thực khuẩn thể (phage) và sự truyền gen
Tái tổ hợp ở vi khuẩn

III.F Sự di truyền ngoài nhân
- Hệ gen ty thể và lạp thể
- Sự di truyền theo dòng mẹ

Sự di truyền theo dòng mẹ đối với quá trình tiến hóa
Hiệu ứng dòng mẹ

III.G Sự di truyền các tính trạng số lượng
- Tính trạng số lượng
- Các đo lường biến dị liên tục
- Các hợp phần phương sai
- Hệ số di truyền theo nghĩa rộng
- Hệ số di truyền theo nghĩa hẹp

Cách đo lường hệ số di truyền
Những hạn chế của hệ số di truyền
Những locut tính trạng số lượng
Những nghiên cứu trên người
Các tính trạng số lượng ở người

III.H Sự xác định giới tính
- Sự xác định giới tính
- Các hệ thống xác định giới tính đơn giản
- Các hệ thống nhiễm sắc thể xác định giới tính
- Sự xác định giới tính ở *Drosophila*

Sự xác định giới tính ở người
Sự xác định giới tính ở thực vật
Các đặc điểm giới tính thứ cấp
Tiến hóa của nhiễm sắc thể giới tính

III.I Sự di truyền liên kết giới tính
- Sự di truyền liên kết giới tính

Các tính trạng bị giới hạn hoặc bị ảnh hưởng bởi giới tính

III.K Lai tế bào soma
- Nhóm liên kết và nhiễm sắc thể
- Nuôi cấy mô và sự dung hợp tế bào

Sắp xếp gen trên các nhiễm sắc thể
Xây dựng bản đồ gen chi tiết

III.L Nội phối
- Nội phối
- Hiệu ứng đối với tần số alen
- Hệ số đồng huyết

Hệ số nội phối
Tầm quan trọng đối với y học

III.M Xác suất
- Xác suất
- Nguyên tắc cộng xác suất
- Nguyên tắc nhân xác suất

Tính xác suất
Phép hoán vị

III.N Phương pháp χ^2 và kiểm tra giả thiết
- Kiểm tra giả thiết
- Phép thử χ^2
- Bậc tự do
- Sử dụng bảng xác suất χ^2

Mức ý nghĩa thống kê
Phép sửa sai số Yates
Phép thử chính xác Fisher

III. A DI TRUYỀN HỌC MENDEL

Các khái niệm cơ bản

Những bằng chứng đầu tiên chứng tỏ sự tồn tại của cái mà ngày nay chúng ta gọi là "gen" bắt nguồn từ công trình nghiên cứu của George Mendel, người đã tiến hành thí nghiệm về đặc điểm di truyền ở cây đậu Hà Lan vào giữa thế kỷ thứ XIX. Trước khi xem xét các kết quả nghiên cứu của ông, chúng ta tìm hiểu một số thuật ngữ cơ bản được sử dụng trong nghiên cứu tính di truyền ở sinh vật bậc cao.

Kiểu hình

Kiểu hình là thuật ngữ được dùng để mô tả mọi đặc điểm (tính trạng) của cơ thể có thể di truyền được, chẳng hạn như màu mắt, hình dạng lá cây, hoặc các bệnh di truyền (ví dụ: bệnh xơ nang). Một con ruồi có thể được mô tả là có kiểu hình mắt đỏ, hoặc một đứa trẻ được gọi là có kiểu hình bị bệnh xơ nang.

Kiểu gen

Là tập hợp tất cả các gen quy định các kiểu hình cụ thể ở một cá thể sinh vật. Trong nghiên cứu di truyền học, ta thường đề cập đến kiểu gen quy định một tính trạng cụ thể. Ví dụ, ruồi giấm thân xám có thể có kiểu gen *BB* hoặc *Bb*.

Dòng thuần

Khái niệm này dùng để chỉ những cơ thể sinh vật thu được từ quá trình nội phối qua nhiều thế hệ, kết quả là một kiểu hình nhất định được duy trì ổn định. Các dòng chó hoặc mèo nòi là những ví dụ cụ thể về các dòng thuần.

Hiện tượng trội

Trong một loài, có thể có những sai khác về kiểu hình đối với một đặc điểm di truyền (chẳng hạn màu mắt có thể là đỏ hoặc trắng). Ở con lai thu được từ phép lai giữa hai cá thể có kiểu hình khác nhau, có thể chỉ có một kiểu hình được biểu hiện, ví dụ: khi lai những ruồi giấm thuần chủng cánh ngắn với ruồi giấm thuần chủng cánh dài, tất cả ruồi con sẽ có cánh dài. Kiểu hình được biểu hiện ở con lai được gọi là kiểu hình trội và kiểu hình kia là kiểu hình lặn. Trong ví dụ trên, cánh dài là trội so với cánh ngắn.

Những yếu tố chủ yếu giúp cho các thí nghiệm của Mendel thành công là việc sử dụng những dòng đậu thuần chủng để tiến hành lai, đồng thời ông đã vận dụng được một số phương pháp phân tích toán học đơn giản để xử lý các kết quả thí nghiệm.

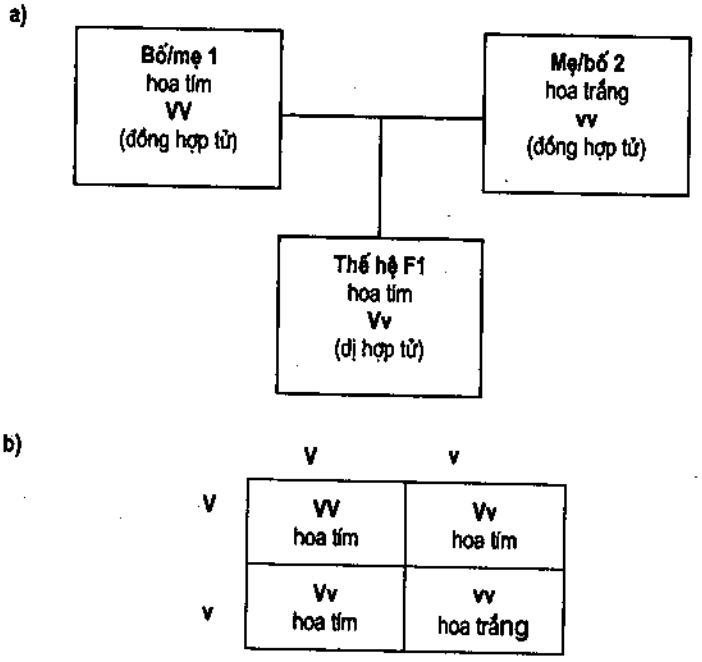
Phép lai đơn tính

Mendel đã nghiên cứu đặc điểm di truyền của một vài kiểu hình ở cây đậu, song chúng ta chỉ phân tích ví dụ ở một kiểu hình, đó là màu sắc cánh hoa. Ông đã tiến hành lai giữa hai dòng thuần, một dòng có cánh hoa màu tím, còn dòng kia có cánh hoa màu trắng. Con lai của phép lai này được ký hiệu là **thế hệ F1** (first filial). Tất cả chúng đều có hoa màu tím. Như vậy, màu tím là trội so với màu trắng. Sau đó ông cho các cây F1 lai với nhau để tạo ra **thế hệ F2** (second filial). Một số cây F2 có hoa màu trắng và những cây khác có hoa màu tím. Tỷ lệ số cây có hoa màu tím so với số cây có hoa màu trắng xấp xỉ tỷ lệ 3:1. Trong các phép lai giữa các cây khác nhau về màu sắc hạt, hình dạng vỏ quả và các kiểu hình khác, Mendel cũng phát hiện thấy tỷ lệ phân ly kiểu hình tương tự. Kiểu hình lặn luôn xuất hiện trở lại ở F2 và luôn chiếm khoảng 1/4 số cây.

Những kết quả thí nghiệm này đã đưa Mendel đến giả thiết rằng tính di truyền là do hoạt động của những nhân tố đặc biệt mà ngày nay chúng ta gọi là gen. Tuy nhiên, kết luận đơn giản này vào thời điểm đó lại trái ngược hẳn với quan điểm thông thường lúc đó cho rằng tính di truyền là do sự trộn lẫn các "dịch di truyền" có nguồn gốc từ bố và mẹ. Rõ ràng chẳng có sự hoà trộn nào xảy ra trong các thí nghiệm của Mendel, cả ở F1, thế hệ mà chỉ có một kiểu hình được biểu hiện, cũng như cả ở F2, thế hệ mà cả hai kiểu hình được biểu hiện riêng rẽ.

Xuất phát từ những kiến thức hiện đại về gen và cấu trúc gen, chúng ta có thể dễ dàng hiểu được cách mà Mendel đã giải thích về những điều đã xảy ra trong thí nghiệm của mình. Điều đó được mô tả vắn tắt bằng sơ đồ trên *Hình 3.1*. Mendel cho rằng, những cây có hoa tím thuần chủng mang hai bản sao của gen quy định sắc tố tím, ký hiệu *V*. Cây hoa trắng mang hai bản sao của dạng biến thể của gen này mã hoá cho hoa màu trắng, ký hiệu là *v*. Chúng ta gọi các cá thể mang hai bản sao giống nhau của một gen là **đồng hợp tử**. Con lai F1 được thừa hưởng hai bản sao khác nhau của gen sắc tố, tức là *Vv*, được gọi là cá thể **dị hợp tử**. Vì alen màu tím (*V*) là trội so với alen màu trắng (*v*) nên tất cả các cây này có hoa màu tím. Khi cho các cá thể F1 lai với nhau, có ba kiểu gen khác nhau được sinh ra ở F2, đó là *VV*, *Vv* và *vv*, với tỷ lệ tương ứng là 1:2:1 (*Bảng 1*). Điều đó đồng nghĩa với tỷ lệ kiểu hình là 3:1 (3 tím : 1 trắng).

Tính đúng đắn của giả thuyết của Mendel càng được củng cố khi cho từng cá thể F2 tự thụ phấn. Tất cả các cây hoa trắng đều thuần chủng. 1/3 số cây hoa màu tím là thuần chủng, còn 2/3 số cây hoa màu tím còn lại sinh ra các cây hoa tím và hoa trắng với tỷ lệ là 3:1. Sau đó, rất nhiều tỷ lệ phân ly 3:1 của các tính trạng do một gen quy định đã được tìm ra ở nhiều loài khác nhau. Một số tính trạng như vậy được liệt kê ở *Bảng 3.2*.



Hình 3.1. **(a) Các giao tử do cây F1 sinh ra mang alen V hoặc v kết hợp ngẫu nhiên, tạo ra thế hệ con có các kiểu gen VV, Vv, vv như được thể hiện ở (b). (b) Mỗi trong số bốn kiểu kết hợp giao tử xuất hiện với tần số bằng nhau. Vì vậy tần số kiểu hình ở F2 là 3 tím: 1 trắng và tần số kiểu gen là 1 VV : 2 Vv : 1 vv.**

Bảng 3.1. **Các sản phẩm do giao phối ở F1**

Kiểu gen	Kiểu hình (màu hoa)	Tỷ lệ
VV	Tím	1
Vv	Tím	2
vv	Trắng	1

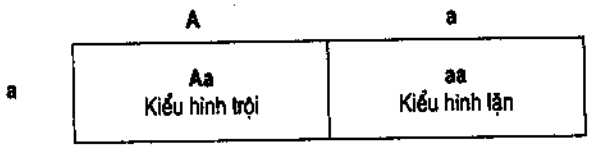
Có một thuật ngữ cũng cần được nêu ở mục này, đó là các gen có thể bị biến đổi gọi là **đột biến**. Các biến thể (trạng thái) khác nhau của một gen được gọi là các **alen**. Cần phân biệt hai thuật ngữ *gen* và *alen*. Ở ví dụ nêu trên, đúng ra ta phải gọi *V* và *v* là hai alen của gen quy định màu sắc cánh hoa. Thông thường, ta ký hiệu alen trội bằng chữ viết hoa và alen lặn bằng chữ viết thường. Tỷ lệ 3:1 được gọi là **tỷ lệ phân ly một tính** (đơn tính) và là tỷ lệ phân ly cơ bản cho tất cả các phương thức di truyền ở sinh vật bậc cao.

Bảng 3.2. **Các ví dụ về sự di truyền do một gen quy định**

Loài	Tính trạng (kiểu hình)
Chuột	Albino/lông bình thường
Cỏ ba lá	Hoa đỏ/hoa trắng
Ruồi giấm	Cánh bình thường/cánh cụt
Người	Mắt xanh/mắt nâu, bệnh xơ nang, bệnh hồng cầu hình liềm, phenylketo niệu

Phát hiện thể dị hợp tử

Tỷ lệ phân ly 1:1 được sinh ra khi lai các cá thể F1 với dạng bố hoặc mẹ đồng hợp tử lặn. Như được trình bày ở *Hình 3.2*, các cá thể F1 dị hợp tử có thể sinh ra hai loại giao tử, mỗi loại chỉ mang hoặc alen trội hoặc alen lặn. Bố hoặc mẹ có kiểu hình lặn chỉ sinh ra một loại giao tử mang alen lặn. Vì thế, đời con có kiểu hình trội và lặn với số lượng tương đương, hay nói cách khác tỷ lệ kiểu hình là 1:1. Phép lai này được gọi là **phép lai phân tích** và được dùng để xác định cá thể dị hợp tử. Đó cũng là kiểu hình kỳ vọng ở những gia đình mà bố, mẹ mang một alen trội hiếm gặp như bệnh Huntington. Vì alen trội này là hiếm nên phần lớn các cá thể mắc bệnh là dị hợp tử.



Hình 3.2. **Tỷ lệ kiểu gen và kiểu hình đều là 1:1 trong phép lai phân tích**

Những tỷ lệ biến đổi của tỷ lệ 3 : 1

Trong nhiều phép lai đơn gen, tỷ lệ phân ly của các kiểu hình có thể không phải là tỷ lệ 3:1. Điều đó có thể do nhiều yếu tố, chẳng hạn những yếu tố sau đây:

Hiện tượng trội một phần hoặc trội không hoàn toàn

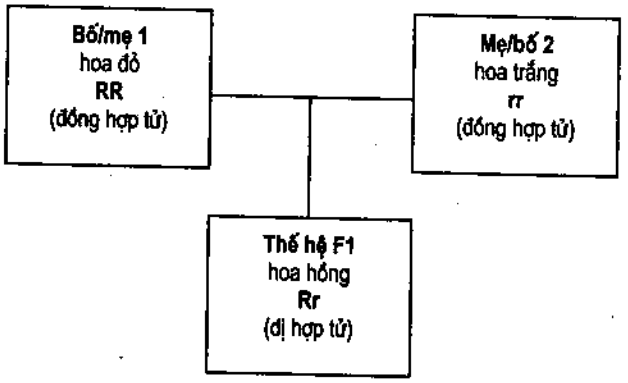
Trong các ví dụ trên đây, tính trạng trội là hoàn toàn. Nghĩa là kiểu hình của các con lai F1 luôn giống với một trong hai kiểu hình của bố, mẹ (kiểu hình trội). Trong trường hợp tính

trạng trội không hoàn toàn (hoặc trội một phần), các cá thể F1 thường có kiểu hình trung gian giữa bố và mẹ. Chẳng hạn như sự di truyền tính trạng màu sắc hoa ở cây hoa mõm chó. Khi tiến hành lai những cây hoa mõm chó có hoa đỏ thuần chủng với những cây có hoa trắng thuần chủng, các con lai ở thế hệ F1 có hoa màu hồng chứ không phải màu đỏ hay màu trắng. Ở thế hệ F2, thu được ba loại hoa theo tỷ lệ phân ly là 1:2:1 (*Bảng 3.3*). Tỷ lệ ba kiểu hình này rõ ràng khác với tỷ lệ 3:1, song rất dễ nhận thấy chúng liên quan với nhau. Phép lai được trình bày trên *Hình 3.3* cho thấy: kiểu gen của những cây hoa đỏ là *RR* và của những cây hoa trắng là *rr*. Ở F1, những cá thể dị hợp tử có kiểu gen là *Rr*. Ở F2, tỷ lệ kiểu hình thu được giống như đã được mô tả ở phần trước: Sự khác biệt ở đây chỉ là các thể dị hợp tử *Rr* có hoa màu hồng nên làm thay đổi tỷ lệ kiểu hình chung của các cá thể ở F2.

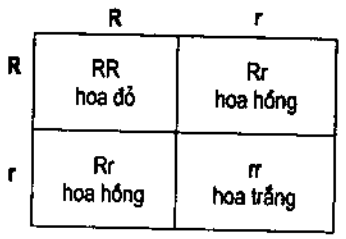
Bảng 3.3. **Sự di truyền màu sắc hoa ở cây Hoa mõm chó**

Kiểu gen	Kiểu hình	Tỷ lệ
rr	trắng	1
Rr	hồng	2
RR	đỏ	1

a)



b)



Hình 3.3. (a) **Các giao tử do cây F1 sinh ra kết hợp ngẫu nhiên tạo ra đời con có các kiểu gen RR, Rr và rr như ở (b). (b) Tỷ lệ kiểu gen giống như ở hình 3.1, song tỷ lệ kiểu hình màu cánh hoa là 1 đỏ : 2 hồng : 1 trắng.**

Các lớp kiểu hình hoa đỏ và hoa trắng ở F2 là đồng hợp tử, do đó tất cả chúng là thuần chủng. Trong khi đó, những cây hoa hồng khi tự thụ phấn luôn sinh ra tỷ lệ kiểu hình ở thế hệ sau là: 1 đỏ : 2 hồng : 1 trắng.

Hiện tượng đồng trội cũng cho kết quả phân ly kiểu hình giống như trường hợp trội không hoàn toàn. Nhưng, đặc điểm phân biệt là ở chỗ: trong trường hợp này, cá thể dị hợp tử

có sự biểu hiện tính trạng của cả hai alen. Chẳng hạn, đặc điểm di truyền các kiểu nhóm máu ở người. Nhóm máu MN do một gen, gồm hai alen là *M* và *N* quy định. Đứa trẻ có bố đồng hợp tử *NN* (có nhóm máu N) và mẹ là đồng hợp tử *MM* (có nhóm máu M) sẽ là dị hợp tử *MN* và có nhóm máu MN. Cả hai kiểu hình đều được biểu hiện ở đời con. Hiện tượng đồng trội cũng làm biến đổi tỷ lệ phân ly kiểu hình 3:1 thành 1:2:1. Các alen có thể được phân biệt bằng các phương pháp phân tử như phản ứng PCR hoặc thẩm tách Southern (Southern botting, xem Mục VI.A và VI.C), lúc đó chúng biểu hiện như những alen đồng trội (cả hai alen đều biểu hiện).

Alen gây chết

Một số alen ảnh hưởng đến sức sống của các cá thể mang nó. Ở hầu hết các trường hợp, các cá thể đồng hợp tử về alen lặn gây chết không sống được nhưng cá thể dị hợp tử có thể sống bình thường. Để phát hiện được cá thể dị hợp tử, người ta cần dựa vào đặc điểm phân ly của kiểu hình nào đó. Chẳng hạn như sự di truyền tính trạng màu lông vàng ở chuột (do alen B^y quy định). Dạng lông vàng có thể xuất hiện ở những dòng chuột có nhiều màu lông khác nhau. Màu lông vàng là trội so với màu lông đen. Chuột có kiểu gen *BB* có lông đen, B^yB^y có lông vàng. Khi cho hai chuột lông vàng giao phối thì tỷ lệ phân ly ở đời con thu được như được nêu trên *Bảng 3.4*. Tuy vậy, do kiểu gen B^yB^y thường gây chết và chuột có kiểu gen này chết trong dạ con. Vì thế, chuột sơ sinh từ phép lai này sẽ có tỷ lệ như ở *Bảng 3.5*. Lúc này, tỷ lệ 3:1 được biến đổi thành tỷ lệ 2:1.

Ở đây có một khái niệm đôi khi có thể gây nhầm lẫn. Alen B^y là lặn khi xét về hiệu quả đối với sức sống, song lại là trội khi xét về tính trạng màu lông. Như vậy, trạng thái trội, lặn của một alen cũng có thể thay đổi tùy thuộc vào kiểu hình phân tích. Trường hợp một gen tham gia quy định nhiều kiểu hình khác nhau là khá phổ biến. Trong trường hợp đó, một alen có thể là trội đối với kiểu hình này, nhưng lại là lặn đối với kiểu hình kia. Một số ví dụ khác cho thấy có các alen gây chết chỉ khi ở trạng thái đồng hợp tử, nhưng lại quy định một tính trạng trội khi ở trạng thái dị hợp tử, chẳng hạn như alen quy định tính trạng không đuôi ở mèo Manx, hay tính trạng chân ngắn ở gà Creeper. Các gen tham gia vào các quá trình sinh trưởng và phát triển của sinh vật thường có các alen gây chết. Đối với những gen này, sự có mặt của một alen đột biến có thể làm thay đổi quá trình sinh trưởng và phát triển bình thường và gây ra những đặc điểm bị biến đổi khác thường ở động vật. Nhưng khi có mặt hai alen đột biến đồng thời, thì mức độ biến đổi đó lớn đến mức có thể gây chết. Hậu quả gây chết có thể xảy ra từ rất sớm, ngay từ khi hợp tử đang phát triển trong dạ con (như ví dụ nêu trên), hoặc ảnh hưởng thông qua việc rút ngắn tuổi thọ của cá thể, chẳng hạn như một số bệnh di truyền ở người, gồm bệnh Tay-Sachs, bệnh Huntigton và bệnh hồng cầu hình liềm.

Ở một số ví dụ khác, phép lai vẫn có thể sinh ra các thể đồng hợp tử song số lượng của nhóm cá thể này giảm đi. Chẳng hạn như tính trạng cánh cụt ở ruồi giấm. Tính trạng do gen lặn (*vg*) quy định. Các alen có hiệu ứng như vậy được gọi là các **alen gây chết một phần**.

Những ví dụ về những tỷ lệ biến đổi từ tỷ lệ phân ly cơ bản là 3:1 lúc đầu dường như khá phức tạp. Song, điều quan trọng là chúng ta phải nhận ra rằng: trong tất cả các trường hợp này, tập tính vận động của gen vẫn giống như trường hợp phân ly theo tỷ lệ cơ bản 3:1, nhưng điểm khác chỉ là sự biến đổi của tỷ lệ kiểu hình. Sự biểu hiện kiểu hình phụ thuộc vào cách thức hoạt động (biểu hiện) của gen thông qua quá trình tổng hợp protein mà gen đó mã hóa, đồng

thời phụ thuộc vào sự tương tác của nhiều loại protein với nhau và với các đại phân tử khác trong tế bào và mô của sinh vật. Chúng ta đã đề cập tới một số nội dung liên quan đến vấn đề này ở phần A. Trong một số phần sau của tài liệu này, chúng ta sẽ đề cập về một số ví dụ cho thấy sự tương tác của các sản phẩm gen sau khi chúng được tổng hợp. Tuy vậy, nếu chúng ta chỉ tập trung để ý đến kiểu gen và sự vận động (di truyền) của các alen từ thế hệ này sang thế hệ khác thì chúng ta dễ dàng nhận ra bằng cơ chế này những gen nhất định quy định sự di truyền của các tính trạng tương ứng.

Bảng 3.4. **Sự di truyền màu lông vàng ở chuột theo lý thuyết trong phép lai BB^y x BB^y**

Kiểu gen	Kiểu hình	Tỷ lệ
BB	Lông đen	1
BB^y	Vàng	2
B^yB^y	Vàng	1

Bảng 3.5. **Sự di truyền thực tế màu lông vàng ở chuột trong phép lai trên**

Kiểu gen	Kiểu hình	Tỷ lệ
BB	Lông đen	1
BB^y	Vàng	2

Đa alen

Các ví dụ nêu trên đây chỉ đề cập đến các gen gồm hai alen. Tuy vậy, trong thực tế có rất nhiều gen tồn tại với nhiều alen khác nhau. Ví dụ gen β-globin, với đột biến ở bộ ba số 6 sinh ra alen quy định bệnh hồng cầu hình liềm, trong khi các đột biến ở những vị trí khác trong gen lại sinh ra hội chứng khác, như β-thalasemia (thiếu máu vùng biển). Cho dù chúng có điểm chung là các trạng thái khác nhau của cùng một gen, nhưng những biến đổi này xảy ra ở những bộ ba mã hóa khác nhau trong gen. Các protein được tạo ra từ các alen biến đổi này có chuỗi β-globin chứa các trình tự amino acid biến đổi tương ứng, vì thế chúng biểu hiện kiểu hình khác nhau.

Ở thỏ, người ta tìm thấy có dãy đa alen của một gen quy định màu lông. Có bốn alen khác nhau trong dãy này, gọi là Agouti, Chinchilla, Himalyan và Albino (bạch tạng). Khi ở trạng thái đồng hợp tử, mỗi một alen tạo ra một màu lông. Khi ở trạng thái dị hợp tử, có hiện tượng trội rõ ràng theo quy luật: Agouti trội so với tất cả các alen còn lại, Chinchilla trội so với Himalyan và Albino, trong khi Himalyan chỉ trội so với Albino. Alen Albino không sinh ra một loại sắc tố nào, vì vậy nó biểu hiện như alen lặn so với tất cả các alen khác.

Một ví dụ khác về dãy đa alen là gen quy định nhóm máu ABO ở người. Chỉ có một gen duy nhất quy định enzym làm nhiệm vụ bổ sung gốc đường vào gốc glycoprotein đặc hiệu trên màng tế bào hồng cầu, nhưng gen này có ba alen. Alen thứ nhất mã hóa cho một dạng enzym bổ sung một phân tử N-acetyl-galactosamine vào glycoprotein, sinh ra nhóm máu A. Alen thứ hai mã hoá cho một biến thể của enzym nêu trên bổ sung galactose thay cho N-acetyl-galactosamine, sinh ra nhóm máu B. Alen thứ ba mã hoá cho một dạng enzym không hoạt động chức năng, không thể thêm gốc đường vào glycoprotein, sinh ra nhóm máu O. Cả ba alen này đều được sinh ra do đột biến từ một gen ban đầu.

Phức hệ hòa hợp mô MHC (*m*ajor *h*istocompability *c*omplex) xác định sự phù hợp của các sơ quan cấy ghép là một ví dụ về một hệ thống đa alen phức hợp. Ngoài những ví dụ này, còn nhiều ví dụ khác về sự tồn tại của các dãy đa alen. Trong một số trường hợp, có thể phân biệt được các alen trực tiếp bằng kỹ thuật phân tích chỉ thị phân tử.

III.B DI TRUYỀN HỌC MENDEL MỞ RỘNG

Phép lai hai tính

Tỷ lệ 3:1 giải thích cho sự di truyền kiểu hình do chỉ một gen duy nhất quy định. Tỷ lệ này gọi là tỷ lệ Mendel cơ bản và những nội dung được đề cập đến sau đây đều liên quan đến tỷ lệ này.

Chúng ta hãy xem xét đồng thời đặc điểm di truyền của hai tính trạng trong phép lai hai tính. Thí nghiệm đơn giản nhất là cho lai hai dòng thuần, một dòng là đồng hợp tử về các alen lặn của cả hai gen và dòng kia là đồng hợp tử gồm hai alen trội của cùng hai gen đó. Mendel tiến hành thí nghiệm đúng như vậy. Trong thí nghiệm của mình, ông dùng các cây đậu khác nhau về hai đặc điểm của hạt.

Việc chọn các đặc điểm của hạt có hai lý do quan trọng:

– Thứ nhất, dễ dàng phân tích chi tiết và xử lý thống kê nhờ số lượng hạt lớn thu được sau mỗi phép lai.

– Thứ hai, rút ngắn thời gian thí nghiệm. Thế hệ F1 là các hạt do kết quả lai các cây bố, mẹ sinh ra, còn thế hệ F2 là các hạt có trong vỏ quả của F1. Như vậy, Mendel tiết kiệm được một năm cho việc gieo trồng các hạt F2 so với khi chọn các cặp tính trạng khác.

Hình dạng hạt do một gen có hai alen quy định, trong đó alen ký hiệu *R* quy định tính trạng hạt trơn, còn *r* là hạt nhăn. Hạt có thể có màu vàng hoặc màu xanh. Đặc điểm này cũng do một gen gồm hai alen quy định. Alen quy định màu vàng, ký hiệu *Y*, là trội hoàn toàn so với alen *y* quy định hạt màu xanh. Thí nghiệm được mô tả như trình bày trên *Hình 3.4*. Phép lai giữa hai dòng bố, mẹ thuần chủng sinh ra các cá thể F1 gồm các hạt đều có màu vàng, trơn. So với phép lai một tính thì F2 có sự đa dạng đáng kể. Có bốn kiểu hình khác nhau. Trong số 556 hạt được phân tích thì có 315 hạt vàng, trơn; 108 hạt xanh, trơn; 101 hạt vàng, nhăn và 32 hạt xanh, nhăn. Sự phân ly này gần với tỷ lệ 9:3:3:1 và được gọi là tỷ lệ phân ly hai tính. Mendel đã thu được các tỷ lệ phân ly gần với tỷ lệ này đối với nhiều tổ hợp các cặp gen và sau đó tỷ lệ này đã được thấy ở nhiều phép lai được tiến hành trên động, thực vật và nấm.

Tỷ lệ mong đợi về kiểu gen và kiểu hình có thể xác định trên sơ đồ bằng cách sử dụng khung Punnett. Ví dụ đơn giản về khung Punnett đã được trình bày ở mục III.A (*Hình 3.1b*) chỉ ra các lớp kiểu gen và kiểu hình có thể thu được từ phép lai đơn tính. Khung Punnett mô tả phép lai hai tính trạng được trình bày trên *Hình 3.4*. Tỷ lệ 9:3:3:1 chỉ đơn giản là sự tổ hợp của hai tỷ lệ 3:1 độc lập và phản ánh các alen của hai gen phân ly độc lập với nhau. Điều đó dễ dàng nhận ra khi xem xét riêng biệt từng gen (*Hình 3.5*). Trước tiên, đối với gen quy định hình dạng hạt: ở F2, tỷ lệ mong đợi là 3 trơn : 1 nhăn. Đối với gen quy định màu sắc hạt: các hạt có tỷ lệ phân ly kiểu hình là 3 vàng : 1 xanh. Nếu nhân tỷ lệ phân ly của hai kiểu hình, ta thu được tỷ lệ phân ly hai tính trạng là 9:3:3:1.

Như đã biết, tỷ lệ phân ly kiểu hình 3:1 có thể bị sai lệch bởi các yếu tố như hiện tượng trội không hoàn toàn hoặc hiệu ứng gây chết của những alen nhất định. Những yếu tố đó cũng

ảnh hưởng tương ứng đến tỷ lệ 9:3:3:1. Ngoài ra, còn có những yếu tố khác nữa cũng có thể ảnh hưởng đến tỷ lệ này. Trong thực tế, tỷ lệ 9:3:3:1 chỉ nghiệm đúng với hai điều kiện:

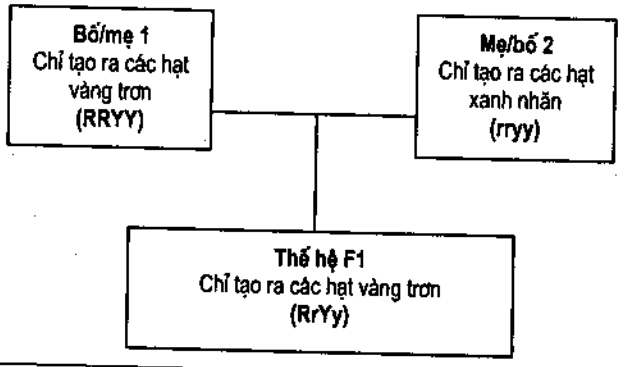
– Thứ nhất, hai gen phải không đồng thời tác động đến cùng một tính trạng. Ví dụ, nếu các protein do hai gen mã hoá cùng tham gia vào một con đường sinh hóa thì tỷ lệ kiểu hình do các kiểu gen sinh ra sẽ thay đổi ở F2. Vấn đề này sẽ được thảo luận chi tiết ở phần sau của mục này (phần tương tác át chế).

– Thứ hai, nếu hai gen liên kết với nhau trên cùng một nhiễm sắc thể thì bốn nhóm giao tử do kiểu gen đó sinh ra sẽ có tần số không bằng nhau. Đó là cơ sở của những nghiên cứu lập bản đồ gen và sẽ được thảo luận chi tiết ở phần liên kết gen.

Cũng giống như trong phép lai đơn tính, người ta có thể tiến hành phép lai phân tích ở con lai F1 trong trường hợp phép lai hai tính. Ví dụ, trong phép lai về màu sắc và hình dạng hạt cho F2 với tỷ lệ phân ly 9:3:3:1 nêu trên, nếu lai con lai F1 với dạng đồng hợp tử lặn về cả hai gen thì đời con sẽ có bốn nhóm kiểu hình: vàng trơn, vàng nhăn, xanh trơn và xanh nhăn. Các nhóm kiểu hình này có tần số bằng nhau (tỷ lệ 1:1:1:1). Phép lai đơn tính có ý nghĩa rất quan trọng trong các nghiên cứu lập bản đồ gen và liên kết gen.

Trong trường hợp nghiên cứu đồng thời ba gen (phép lai ba tính), trong đó các alen của cả ba gen đều biểu hiện trội hoàn toàn, thì tỷ lệ phân ly kiểu hình sẽ là 27:9:9:9:3:3:3:1. Có thể xác định tỷ lệ phân ly kiểu hình mong đợi cho các phép lai với số gen bất kỳ bằng biểu thức $(3:1)^n$, trong đó n là số gen.

a)



b)

Giao tử	RY	Ry	rY	ry
RY	RRYY Vàng trơn	RRYy Vàng trơn	RrYY Vàng trơn	RrYy Vàng trơn
Ry	RRYy Vàng trơn	Rryy Xanh trơn	RrYy Vàng trơn	Rryy Xanh trơn
rY	RrYY Vàng trơn	RrYy Vàng trơn	rrYY Vàng nhăn	rrYy Vàng nhăn
ry	RrYy Vàng trơn	Rryy Xanh trơn	rrYY Vàng nhăn	rryy Xanh nhăn

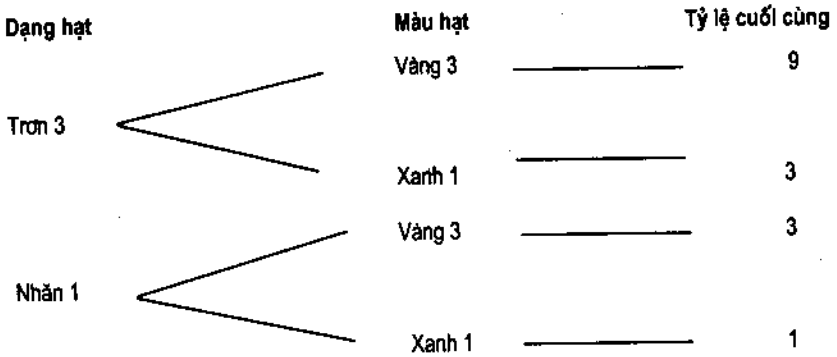
Các lớp kiểu hình
Hạt vàng, trơn
Hạt xanh, trơn
Hạt vàng, nhăn
Hạt xanh, nhăn

Các lớp kiểu gen
RRYY (1), RRYy (2), RrYY (2), RrYy (4)
Rryy (1), Rryy (2)
rrYY (1), rrYy (2)
rryy (1)

Hình 3.4. **(a) Sự sản sinh ra các cây F1. Các cây F1 có thể tạo ra bốn loại giao tử khác nhau: RY, Ry, rY và ry. (b) Khung Punnett cho thấy các kiểu gen có thể có và các kiểu hình xuất hiện khi các cây F1 giao phối với nhau. Các con số trong ngoặc là số lần xuất hiện của kiểu gen trong khung Punnett.**

Tương tác át chế

Tương tác át chế là trường hợp không thu được tỷ lệ kiểu hình mong đợi do hình thức tương tác sinh lý nào đó xảy ra giữa các gen cùng tham gia quy định tính trạng. Hiện tượng này thường biểu hiện thành các tỷ lệ biến đổi từ tỷ lệ phân ly 9:3:3:1 với sự giảm số lượng các nhóm kiểu hình khác nhau. Có nhiều tỷ lệ khác nhau bắt nguồn từ tỷ lệ 9:3:3:1. Ba tỷ lệ trong số đó được trình bày ở *Hình 3.6, 3.7 và 3.8*. Mặc dù mỗi trường hợp có thể được giải thích bằng những ví dụ riêng biệt lấy từ các tài liệu tham khảo, song sẽ đơn giản hơn nếu các tỷ lệ đó được mô tả thông qua những con đường sinh hoá mô phỏng ở một loài cây không tên, trong đó có hai enzym được mã hoá bởi hai gen riêng biệt cùng tham gia vào việc sản sinh sắc tố cánh hoa. Các gen được ký hiệu bằng các chữ in hoa *A* và *B* là các alen trội hoạt động chức năng và các alen được ký hiệu bằng các chữ in thường là các alen lặn không hoạt động chức năng. Tất cả các ví dụ đều có F2 được sinh ra từ F1 dị hợp tử kép *AaBb*. Sau khi xây dựng sơ đồ lai, lập khung Punnett, chúng ta xác định được sự phân ly tỷ lệ các kiểu hình như nêu trên *Hình 3.4*.

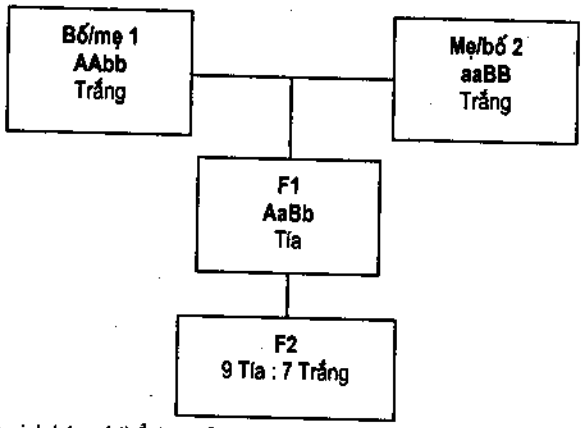


Hình 3.5. **Sự phân ly theo tỷ lệ 9:3:3:1 (xem chi tiết trong phần diễn giải)**

Theo nghĩa hẹp trước đây, tỷ lệ 12:3:1 là tỷ lệ duy nhất được coi là do kết quả tương tác át chế của các alen, lúc này alen *A* hoàn toàn ức chế biểu hiện của alen *B*. Tuy nhiên, thuật ngữ tương tác át chế giờ đây được mở rộng ra đối với mọi trường hợp có sự tương tác giữa các gen dẫn đến sự thay đổi tỷ lệ mong đợi.

Có một số tỷ lệ biến đổi khác so với tỷ lệ 9:3:3:1 là do sự tương tác giữa các sản phẩm của gen tạo ra. Những tỷ lệ này bao gồm 9:6:1, 15:1 và 13:3. Chúng ta có thể nhận thấy những tỷ lệ này có thể thu được khi xét các con đường sinh hoá mà có sự tham gia của các sản phẩm gen. Nhớ rằng, trong mỗi trường hợp, tỷ lệ biến đổi đó đều bắt nguồn từ cách cộng một vài lớp kiểu hình trong số bốn lớp kiểu hình từ tỷ lệ cơ bản là 9:3:3:1.

Các ví dụ mô tả ở đây là những ví dụ đơn giản đã được chọn lọc nhưng nó minh hoạ được các nguyên lý cơ bản nhất. Chúng ta biết rằng khái niệm kiểu hình có khả năng được hiểu theo những các cách khác nhau. Nhà chọn giống thực vật có thể chỉ sử dụng màu sắc hoa như chúng ta đã sử dụng để mô tả kiểu hình, nhưng nhà sinh hoá thực vật có thể lại xem xét các kiểu hình dưới góc độ các quá trình sinh hóa, như việc phân tích các sản phẩm của gen như các enzym A và B trong điều kiện *invitro* chẳng hạn. Tỷ lệ kiểu hình được xác định theo cách đó có thể khác với các tỷ lệ đã nêu. Tuy vậy, trong mọi trường hợp, tỷ lệ kiểu gen vẫn không thay đổi.



Trong ví dụ này, con đường sinh hóa có thể đơn giản là enzym A chuyển hóa cơ chất của nó thành sản phẩm màu trắng và là cơ chất cho enzym B chuyển hoá tiếp thành sản phẩm màu tía

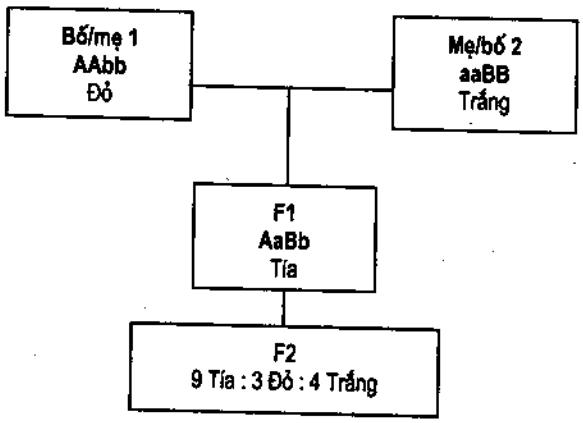
Cơ chất trắng —(Enzym A)→ Sản phẩm trắng —(Enzym B)→ Sản phẩm tía

Các kiểu hình và kiểu gen F2

Tía: 9	
AABB (1), AABb (2), AaBB (2), AaBb (4)	Phải có ít nhất một bản sao của alen A và alen B, nên tổng hợp được sắc tố tía
Trắng: 7	
AAbb (1), Aabb (2)	Chỉ có enzym A nên không hoàn chỉnh để có thể tổng hợp sắc tố tía
aaBB (2), aaBb (1)	Chỉ có enzym B nên không hoàn chỉnh để có thể tổng hợp sắc tố tía
aabb (1)	Không có enzym nào nên không tổng hợp được sắc tố tía

Trường hợp này gọi là tương tác bổ trợ (complementation), do hai alen trong hai dòng màu trắng có thể bổ trợ cho nhau tạo ra sản phẩm màu tía ở con lai F1.

Hình 3.6. **Tỷ lệ phân ly 9 : 7**



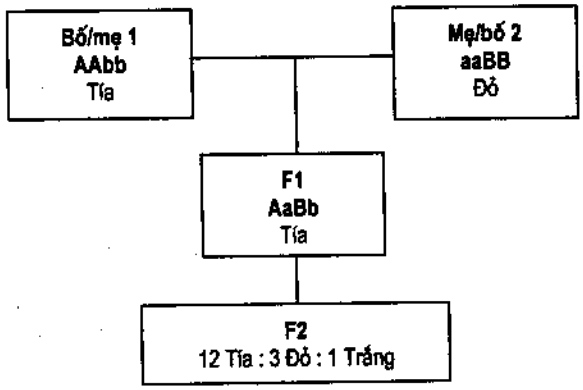
Trong ví dụ này, con đường sinh hóa cũng đơn giản, nhưng sản phẩm của enzym A có màu đỏ

Cơ chất trắng —(Enzym A)→ Sản phẩm đỏ —(Enzym B)→ Sản phẩm tía

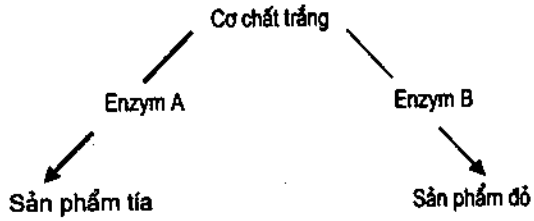
Các kiểu hình và kiểu gen F2

Tía: 9	
AABB (1), AABb (2), AaBB (2), AaBb (4)	Phải có ít nhất một bản sao của cả hai alen A và B, nên tổng hợp được sắc tố tía
Đỏ: 3	
AAbb (1), Aabb (2)	Chỉ có enzym A nên tạo được sắc tố đỏ, nhưng sắc tố đỏ không chuyển hóa được thành sắc tố tím
Trắng: 4	
aaBB (2), aaBb (1)	Không có enzym A nên không thể tổng hợp được sắc tố đỏ là cơ chất của enzym B, nên duy trì màu trắng
aabb (1)	Không có enzym nào nên không tổng hợp được sắc tố

Hình 3.7. **Tỷ lệ phân ly 9 : 3 : 4.**



Trong trường hợp này, con đường sinh hóa phức tạp hơn. Ở đây, hai enzym sử dụng cùng cơ chất. Enzym A chuyển hóa cơ chất thành sản phẩm có màu tía có ái lực với cơ chất cao hơn so với enzym B chuyển hóa cơ chất thành sản phẩm có màu đỏ. Sự khác biệt về ái lực đối với cơ chất rõ rệt đến mức enzym B chỉ có thể hoạt động tốt khi không có mặt enzym A.



Các kiểu hình và kiểu gen F2	
Tía: 12	
AABB (1), AABb (2), AaBB (2), AaBb (4), AAbb (1), Aabb (2)	Những kiểu gen này có alen A và chuyển hóa cơ chất thành sản phẩm có màu tía
Đỏ: 3	
aaBB (2), aaBb (1)	Thiếu enzym A, nhưng có enzym B nên chuyển hóa cơ chất thành chất có màu đỏ
Trắng: 1	
aabb (1)	Không có enzym nào nên không tổng hợp được sắc tố

Hình 3.8. **Tỷ lệ phân ly 12 : 3 : 1.**

Các quy luật Mendel

Các tỷ lệ phân ly trong các phép lai một tính và hai tính xuất phát trực tiếp từ các thí nghiệm của Mendel. Chúng ta nên nhớ rằng, những thí nghiệm của Mendel được thực hiện khi chưa rõ khái niệm về nhiễm sắc thể và ADN. Công trình của Mendel thường được diễn đạt thành hai quy luật hay còn gọi là hai nguyên lý mà ông đã dùng để giải thích quá trình di truyền của các kiểu hình.

Quy luật thứ nhất phát biểu rằng: trong quá trình hình thành giao tử, hai alen của cùng một gen phân ly (tách nhau ra) sao cho mỗi giao tử chỉ nhận một alen. Điều này chứng minh rõ ràng cho tỷ lệ phân ly đơn tính cơ bản là 3:1.

Quy luật thứ hai phát biểu rằng: các alen của bất kỳ gen nào cũng phân ly độc lập với các alen của gen khác. Điều đó bắt nguồn từ các số liệu của phép lai hai tính được mô tả ở *Hình 3.4*. Bốn alen *R*, *r*, *Y* và *y* phải hoạt động độc lập với nhau nên mới có bốn loại giao tử *RY*, *Ry*, *rY* và *ry* xuất hiện với số lượng bằng nhau. Hai nguyên lý này là cơ sở của những nguyên lý vận động trong quá trình di truyền của các gen.

Kiểm tra số liệu thực nghiệm

Cả sinh viên khi làm thí nghiệm cũng như các nhà di truyền học khi nghiên cứu đều phải xử lý các số liệu thu được ở F1 và F2 từ các phép lai. Người thực hiện thí nghiệm lai phải có khả năng nhận biết các tỷ lệ để xác định số gen trong quá trình nghiên cứu và xem xét có hay không sự tương tác bổ trợ giữa các gen. Có thể bắt gặp một số trường hợp sau:

Trong phép lai giữa những cây cà chua thuần chủng quả màu trắng và những cây quả màu tía, các cây lai F1 đều có quả màu tía. F2 thu được 160 cây, trong đó có 99 cây quả tía, 25 cây quả đỏ và 36 cây quả trắng. Chúng ta phải giải thích kết quả đó như thế nào?

Vì chúng ta không biết gì về các gen kiểm soát màu sắc quả ở cà chua nên trước hết có thể tự hỏi "Các số liệu này có phù hợp với tỷ lệ nào trong số các tỷ lệ của Mendel?". Cách đơn giản nhất là loại bỏ những tỷ lệ rõ ràng không thể áp dụng được. Vì ở F2 có ba kiểu hình khác nhau, nên các tỷ lệ với hai lớp kiểu hình như 9:7 hoặc 3:1 sẽ bị loại. Có thể là tỷ lệ 9:3:4, nhưng cũng có thể áp dụng với tỷ lệ 1:2:1. Vậy, bằng cách nào chúng ta có thể xác định tỷ lệ nào là phù hợp nhất với các số liệu thực nghiệm của chúng ta?

Việc so sánh tỷ lệ thực nghiệm và tỷ lệ lý thuyết (tỷ lệ mong chờ) có thể được thực hiện bằng cách sử dụng phương pháp thống kê χ^2. Phương pháp này được trình bày chi tiết ở mục III.N. Ở đây ta chỉ áp dụng nó. Các số liệu thực nghiệm quan sát được so sánh với các số liệu lý thuyết thu được bằng cách tính theo tỷ lệ phân ly mà chúng ta giả thiết là phù hợp nhất, trong trường hợp tỷ lệ là 9:3:4, ta có:

Số liệu quan sát: 99 tía : 25 đỏ : 36 trắng. Tổng cộng: 160

Số liệu mong đợi tính theo tỷ lệ 9:3:4

9/16 x 160		3/16 x 160		4/16 x 160	
90		30		40	
$\chi^2 = \frac{(99-90)^2}{90}$	+	$\frac{(30-25)^2}{30}$	+	$\frac{(40-36)^2}{40}$	$= 2,31$

Với số bậc tự do bằng 2 cho thấy kết quả thí nghiệm không khác với tỷ lệ 9:3:4 tại xác suất 5%.

Áp dụng phương pháp trên với tỷ lệ 1:2:1, ta có

1/4 x 160		2/4 x 160		1/4 x 160	
40		80		40	
$\chi^2 = \frac{(40-25)^2}{40}$	+	$\frac{(99-80)^2}{80}$	+	$\frac{(40-39)^2}{40}$	$= 10,53$

Với số bậc tự do bằng 2, cho thấy kết quả thu được khác với tỷ lệ 1:2:1 tại mức xác suất 5%.

Trên cơ sở đó, có thể nói tỷ lệ 1:2:1 bị loại và tỷ lệ 9:3:4 là phù hợp.

III.C GIẢM PHÂN VÀ QUÁ TRÌNH PHÁT SINH GIAO TỬ

Sự sinh sản

Sự sinh sản được thực hiện nhờ một trong hai phương pháp, vô tính hoặc hữu tính. Sự sinh sản vô tính là sự sinh ra cá thể mới từ những tế bào hoặc mô của một sinh vật. Quá trình này phổ biến ở thực vật và nhiều vi sinh vật. Nó có thể được thực hiện nhờ sự phân đôi các vi khuẩn đơn bào hoặc sinh ra từ các bào tử vô tính chuyên hoá. Các quá trình sinh sản vô tính đã và đang được sử dụng cho mục đích thương mại, chẳng hạn như nhân giống vô tính ở thực vật. Đến nay, người ta đã có thể tái sinh toàn bộ cơ thể từ một số loại tế bào đã chuyên hóa. Điều này đầu tiên được thực hiện thành công ở cây cà rốt và ếch; sau này đã được thực hiện thành công ở cừu, chuột và một số loài động vật có vú khác. Các sinh vật sinh sản vô tính về nguyên tắc có vật chất di truyền giống hệt nhau và giống với bố, mẹ chúng. Nhóm tập hợp các cá thể sinh vật giống hệt nhau về vật chất di truyền như vậy được gọi là dòng vô tính.

Sinh sản hữu tính khác ở chỗ có sự kết hợp của các tế bào (giao tử) có nguồn gốc từ các cá thể bố, mẹ khác nhau để hình thành hợp tử. Các quá trình di truyền tham gia vào sự hình thành giao tử cho phép có những biến đổi di truyền ở thế hệ con. Sự sinh sản hữu tính chỉ hạn chế ở những loài sinh vật lưỡng bội hoặc ở những loài có một giai đoạn lưỡng bội trong vòng đời của chúng.

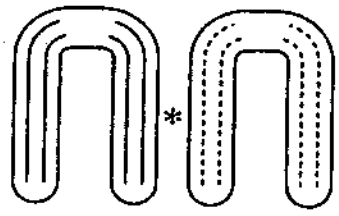
Sự sản sinh các giao tử được gọi là quá trình phát sinh giao tử. Đó có thể là một quá trình phức tạp bao gồm sự biệt hoá giới tính và sản sinh ra các giao tử đực và cái có mức độ chuyên hoá cao; hoặc ở các sinh vật nhân chuẩn bậc thấp, đó là sự kết hợp của các tế bào giống hệt nhau (gọi là sự đẳng giao). Dù ở đối tượng nào thì sự phát sinh giao tử đều phải làm cho số lượng nhiễm sắc thể (NST) giảm đi một nửa, nếu không thì ở thế hệ kế tiếp, số nhiễm sắc thể sẽ tăng gấp đôi so với bố, mẹ. Trong quá trình phát sinh giao tử, số lượng NST giảm đi một nửa nhờ một kiểu phân chia tế bào đặc biệt, gọi là giảm phân. Trong quá trình giảm số lượng NST đi một nửa, các tổ hợp alen được sắp xếp lại, tạo ra các giao tử tái tổ hợp. Có hai quá trình riêng biệt tham gia vào sự kiện này: sự phân ly độc lập của các NST và sự trao đổi chéo.

Giảm phân

Sự phân chia giảm phân rất khác với nguyên phân. Giảm phân gồm hai lần phân chia tế bào liên tiếp để tạo ra bốn tế bào con, mỗi tế bào có một nửa số lượng NST so với tế bào mẹ. Giữa hai lần phân bào, được gọi lần lượt là giảm phân I và giảm phân II không có sự nhân đôi thêm nữa của ADN.

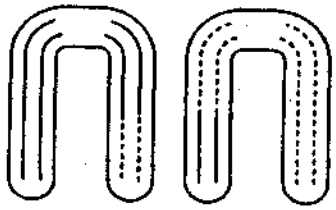
Giống như trong nguyên phân, giảm phân cũng được phân chia thành bốn kỳ, là kỳ đầu, kỳ giữa, kỳ sau và kỳ cuối. Tuy vậy, tập tính của các nhiễm sắc thể ở lần phân bào thứ nhất của giảm phân rất khác so với ở nguyên phân.

Ở kỳ đầu, mỗi NST bắt cặp với chiếc tương đồng của nó. Cấu trúc bắt cặp lúc này của cặp NST tương đồng được gọi là thể lưỡng trị (đôi khi thuật ngữ thể tứ trị cũng được dùng để mô tả các cấu trúc này, nhưng để tránh nhầm lẫn nên dùng thuật ngữ thể tứ trị để chỉ bốn tế bào được tạo ra ở cuối giảm phân). Ngoài sự hình thành phức hệ tiếp hợp, các NST trong cặp tương đồng còn "gắn" với nhau qua các vắt chéo. Đó là các cấu trúc nơi các nhiễm sắc thể tương đồng có thể trao đổi thuận nghịch với nhau các đoạn nhiễm sắc tử.



Hai nhiễm sắc thể tương đồng (trước đó, đã nhân đôi thành các nhiễm sắc tử) kết cặp với nhau. Các chuỗi xoắn kép ADN được biểu diễn bằng các đường ═══ và ===.

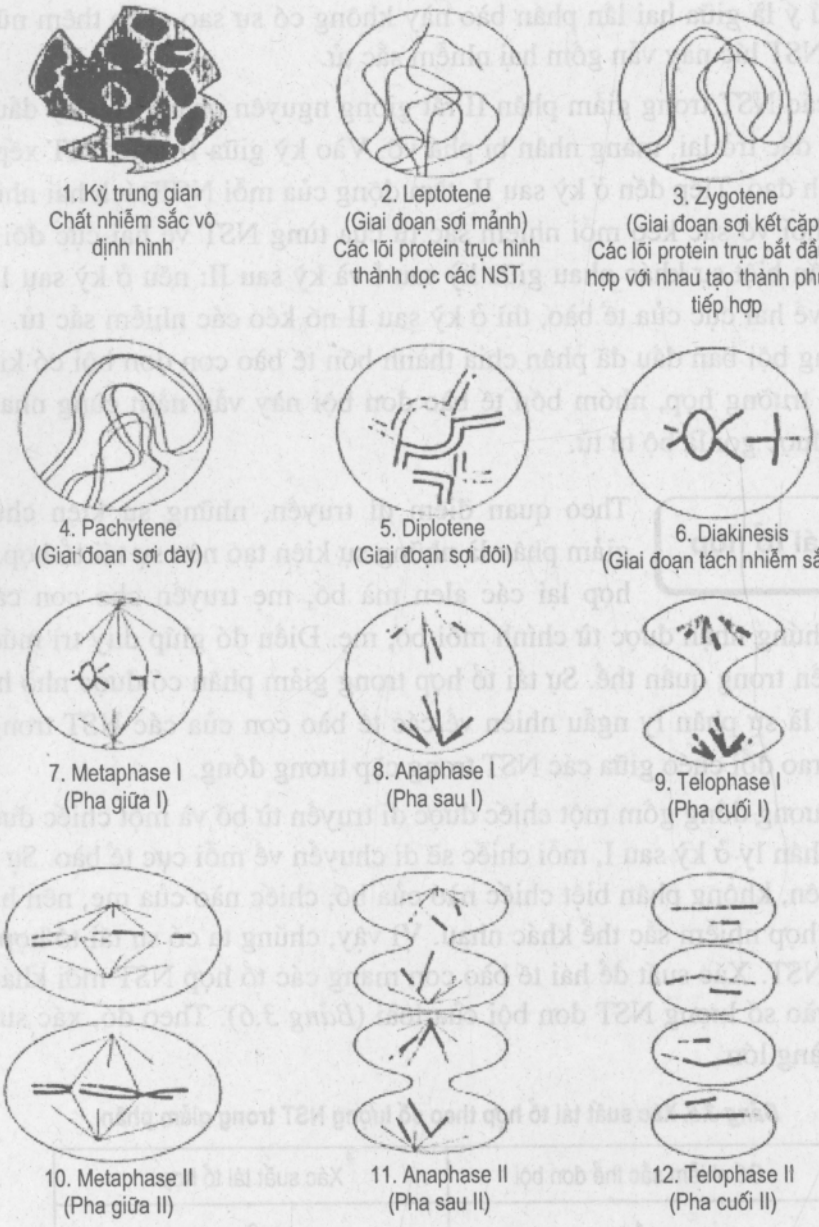
Các nhiễm sắc tử đứt gãy ở vị trí hoa thị (*) và trao đổi tương hỗ các trình tự ADN và các alen tại các locut gen tương ứng trên hai nhiễm sắc tử



Hình 3.9. **Cơ sở phân tử của trao đổi chéo**

Quá trình trao đổi này bao gồm sự đứt gãy chuỗi xoắn kép ở mỗi nhiễm sắc tử tại các điểm tương ứng, rồi sau đó ADN của nhiễm sắc tử này nối lại với ADN của nhiễm sắc tử kia. Sự trao đổi vật chất di truyền này được gọi là sự trao đổi chéo. Cơ sở phân tử của quá trình này được minh họa trên *Hình 3.9*.

Diễn biến tế bào học của giảm phân được mô tả trên *Hình 3.10*. Kỳ đầu của giảm phân I được chia thành năm giai đoạn. Ở giai đoạn đầu, **leptotene** (gọi là "**giai đoạn sợi mảnh**"), chất nhiễm sắc kết lại thành những sợi rất mảnh và dài, xuất hiện như đám rối trong nhân tế bào. Sau đó các nhiễm sắc thể ngày càng trở nên ngắn lại và dày hơn. Ở giai đoạn **zygotene** ("**giai đoạn sợi kết cặp**"), các nhiễm sắc thể tương đồng có cấu trúc ghép đôi từng phần. Tại giai đoạn này, các nhiễm sắc thể vẫn còn dài và các cặp nhiễm sắc thể có thể còn chồng chéo hoặc xoắn lấy nhau cục bộ. Đến giai đoạn thứ ba, **pachytene** ("**giai đoạn sợi dày**"), sự ghép đôi (tiếp hợp) giữa cặp NST tương đồng xảy ra hoàn toàn. Sự tiếp hợp này của các NST tương đồng là một quá trình diễn ra rất chính xác. Cơ chế phân tử của quá trình này còn chưa được hiểu biết hoàn toàn, nhưng các quan sát tế bào học cho thấy có sự hình thành cấu trúc giữa các NST gọi là phức hệ tiếp hợp. Phức hệ này chứa các protein cho phép liên kết giữa các đoạn tương đồng với nhau. Dưới kính hiển vi điện tử, nó có hình dạng như một chiếc khoá kéo "phéc-mơ-tuya". Người ta cho rằng phức hệ tiếp hợp là yếu tố chính trong quá trình ghép đôi của cặp NST tương đồng trong kỳ đầu của giảm phân. Sự chuyển dịch từ giai đoạn pachytene sang **diplotene** ("**giai đoạn sợi đôi**") diễn ra khi các NST tương đồng bắt đầu tách nhau ra (khỏi phức hệ tiếp hợp), nhưng các thể lưỡng trị thường vẫn còn "gắn" với nhau ở các cấu trúc vắt chéo. Lúc này, màng nhân và nhân con dần bị phá vỡ. Sau đó, khi các NST tiếp tục kết đặc thì tế bào chuyển sang giai đoạn **diakinesis** ("**giai đoạn tách nhiễm sắc tử**") là giai đoạn cuối cùng của kỳ đầu giảm phân I. Ở các giai đoạn diplotene và diakinesis, bốn nhiễm sắc tử (các nhiễm sắc thể con) ở các thể lưỡng trị định hình rõ dần dần và có thể nhìn thấy được dưới kính hiển vi các cấu trúc vắt chéo.



Hình 3.10. **Quá trình giảm phân**

Ở kỳ giữa giảm phân I (gọi tắt là kỳ giữa I), các thể lưỡng trị có các tâm động gắn với các sợi thoi vô sắc và dần dịch chuyển về tập trung thành hàng trên mặt phẳng xích đạo của tế bào. Ở kỳ sau I, các sợi thoi vô sắc kéo mỗi NST của các cặp NST tương đồng về các cực đối nhau của tế bào, kéo theo các cánh của nó, dẫn đến sự phân ly đồng đều của hai bộ NST về hai cực của tế bào. Ở kỳ cuối I, màng nhân hình thành và các nhiễm sắc thể giải kết đặc. Ở một số loài, quá trình giảm phân có thể bỏ qua bước này, hoặc chỉ trải qua một giai đoạn kỳ cuối I rất ngắn, rồi chuyển ngay sang lần phân bào thứ hai (giảm phân II). Kết quả của giảm phân I là số lượng NST của tế bào con giảm đi một nửa so với tế bào ban đầu, hay nói cách khác, số lượng lưỡng bội đã giảm xuống thành đơn bội. Giảm phân II nhanh chóng bắt đầu sau kỳ cuối I, với

đặc điểm đáng chú ý là giữa hai lần phân bào này không có sự sao chép thêm nữa của ADN. Thay vào đó, mỗi NST lúc này vẫn gồm hai nhiễm sắc tử.

Tập tính của các NST trong giảm phân II rất giống nguyên phân. Vào kỳ đầu II, ở cả hai nhân, các NST kết đặc trở lại, màng nhân bị phá vỡ. Vào kỳ giữa II, các NST xếp thành hàng trên mặt phẳng xích đạo. Tiếp đến ở kỳ sau II, tâm động của mỗi NST tách hai nhiễm sắc tử ra khỏi nhau và sợi thoi vô sắc kéo mỗi nhiễm sắc tử của từng NST về hai cực đối nhau của tế bào. Cần chú ý phân biệt sự khác nhau giữa kỳ sau I và kỳ sau II: nếu ở kỳ sau I, thoi vô sắc kéo NST trọn vẹn về hai cực của tế bào, thì ở kỳ sau II nó kéo các nhiễm sắc tử. Vào cuối kỳ cuối II, tế bào lưỡng bội ban đầu đã phân chia thành bốn tế bào con đơn bội có kiểu gen khác nhau. Trong nhiều trường hợp, nhóm bốn tế bào đơn bội này vẫn nằm cùng nhau trong một cấu trúc chung và được gọi là bộ tứ tử.

Giảm phân và tái tổ hợp

Theo quan điểm di truyền, những sự kiện chủ yếu trong giảm phân là những sự kiện tạo nên sự tái tổ hợp, tức là sự tổ hợp lại các alen mà bố, mẹ truyền cho con cái, khác với những tổ hợp mà chúng nhận được từ chính mỗi bố, mẹ. Điều đó giúp duy trì mức độ biến dị (tái tổ hợp) di truyền trong quần thể. Sự tái tổ hợp trong giảm phân có được nhờ hai quá trình khác biệt nhau, đó là sự phân ly ngẫu nhiên về các tế bào con của các NST trong cặp tương đồng và khả năng trao đổi chéo giữa các NST trong cặp tương đồng.

Mỗi cặp NST tương đồng gồm một chiếc được di truyền từ bố và một chiếc được di truyền từ mẹ. Khi chúng phân ly ở kỳ sau I, mỗi chiếc sẽ di chuyển về mỗi cực tế bào. Sự phân ly này hoàn toàn ngẫu nhiên, không phân biệt chiếc nào của bố, chiếc nào của mẹ, nên hai nhân đơn bội sẽ mang các tổ hợp nhiễm sắc thể khác nhau. Vì vậy, chúng ta có sự tái tổ hợp do sự phân ly độc lập của các NST. Xác suất để hai tế bào con mang các tổ hợp NST mới khác với tế bào bố, mẹ phụ thuộc vào số lượng NST đơn bội của loài (*Bảng 3.6*). Theo đó, xác suất càng cao khi số lượng NST càng lớn.

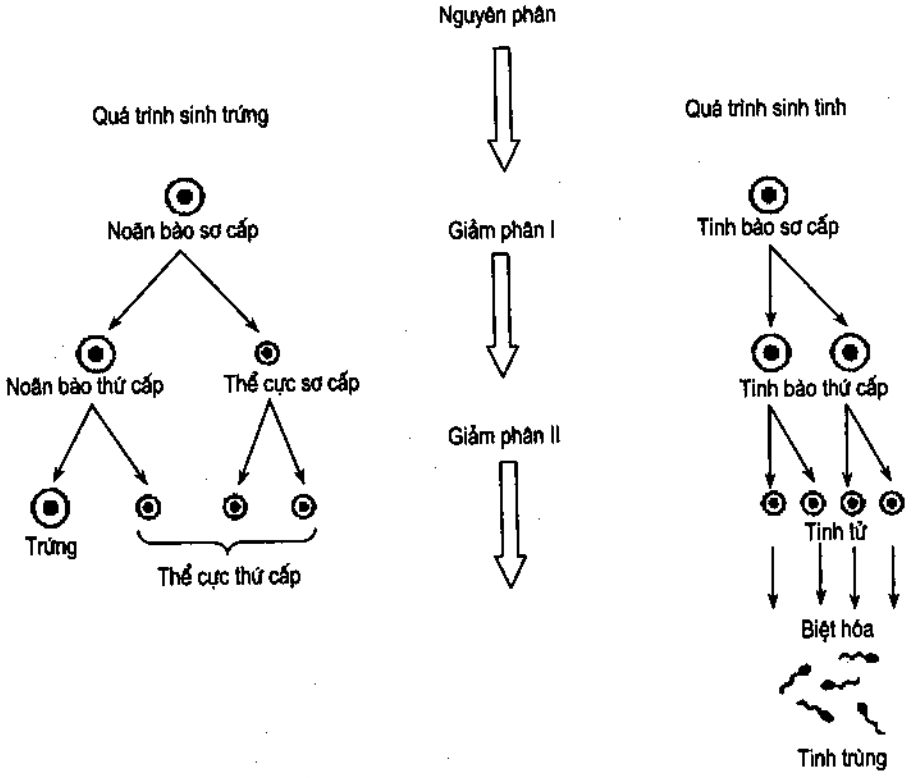
Bảng 3.6. **Xác suất tái tổ hợp theo số lượng NST trong giảm phân**

Số nhiễm sắc thể đơn bội	Xác suất tái tổ hợp
1	0
2	0,5
3	0,75
...	...
n	$(2^n - 2)/2^n$

Tỷ lệ phân ly trong phép lai hai tính của Mendel cho thấy có sự tổ hợp tự do của các gen. Vì bốn loại giao tử được tạo ra với tỷ lệ bằng nhau nên tỷ lệ phân ly thu được là 9:3:3:1. Tần số tái tổ hợp thực tế sẽ còn cao hơn nữa do kết quả của trao đổi chéo. Sự trao đổi các vùng nhiễm sắc tử được di truyền từ bố và mẹ sẽ tạo ra các tổ hợp gen mới được thế hệ kế tiếp truyền lại trên cùng một nhiễm sắc thể. Sự đứt và nối lại chuỗi xoắn kép ADN là cơ chế của sự tái tổ hợp và là cơ chế chính đối với nhiều hiện tượng di truyền xảy ra cả ở sinh vật nhân chuẩn cũng như ở sinh vật nhân sơ.

Sự phát sinh giao tử

Ở động vật, giảm phân chỉ xảy ra ở buồng trứng và tinh hoàn (thậm chí ở các mô này nó cũng chỉ giới hạn ở các tế bào tạo giao tử, được gọi là các dòng tế bào mầm). Dù biểu hiện của quá trình phát sinh giao tử có hơi khác nhau ở các loài sinh vật khác nhau, quá trình này ở các loài sinh vật đều có những bước chung như quá trình phát sinh giao tử ở động vật có vú được mô tả dưới đây (*Hình 3.11*).



Hình 3.11. **Quá trình sinh tinh và sinh trứng ở động vật có vú**

Trong quá trình phát sinh giao tử đực (quá trình sinh tinh trùng), các tế bào tiền thân của các tế bào mầm trải qua nhiều lần nguyên phân để duy trì vốn tinh nguyên bào rồi biệt hoá thành các tinh bào sơ cấp. Giảm phân xảy ra trong các tế bào này. Sau lần giảm phân I, các tế bào này được gọi là các tinh bào thứ cấp. Chúng là các tế bào đơn bội. Sản phẩm của lần giảm phân II là các tinh tử. Các tinh tử này sau đó được biệt hoá thành tinh trùng hoàn thiện, linh động và có đuôi.

Quá trình sinh trứng ở con cái cũng tương tự như vậy. Các noãn nguyên bào cũng trải qua nguyên phân trước khi biệt hoá thành các noãn bào sơ cấp. Sau đó các noãn bào sơ cấp giảm phân. Cả hai tế bào được tạo ra sau giảm phân I đều đơn bội nhưng có kích thước rất khác nhau. Tế bào to được gọi là noãn bào thứ cấp, còn tế bào nhỏ được gọi là thể cực sơ cấp. Hai tế bào này vẫn gắn với nhau. Cả hai đều trải qua giảm phân II. Noãn bào thứ cấp phân chia tế bào chất không đều, tạo ra một tế bào trứng lớn (chứa hầu hết tế bào chất), và một thể cực thứ cấp nhỏ. Thể cực sơ cấp cũng phân chia để tạo ra hai thể cực thứ cấp nữa (ở một số loài có thể không có hiện tượng này). Chỉ các tế bào trứng tham gia truyền gen cho thế hệ sau. Như vậy sự sinh trứng khác với sự sinh tinh ở chỗ chỉ có một trong số bốn sản phẩm của giảm phân là

hình thành giao tử. Một khác biệt khác nữa là thời gian xảy ra giảm phân. Nếu sự sinh tinh thường diễn ra tương đối liên tục kể từ khi cá thể trưởng thành (thành thục về sinh dục), thì trong quá trình sinh trứng, các noãn bào sơ cấp được hoàn thành từ giai đoạn bào thai. Các noãn bào sơ cấp này giữ nguyên ở giai đoạn kỳ đầu giảm phân I cho đến khi trưởng thành. Giảm phân I chỉ thực sự hoàn chỉnh khi trứng rụng, và giảm phân II diễn ra sau thụ tinh. Ở người phụ nữ, thời gian các noãn bào sơ cấp dừng ở giảm phân I có thể kéo dài đến 45 năm. Tuy vậy, sự kéo dài thời gian này có liên quan đến việc tăng tần số sinh trẻ dị bội ở những người mẹ cao tuổi.

Sự phát sinh các giao tử dị bội

Nguyên nhân chính của hiện tượng dị bội thể là sự biểu hiện tập tính sai lầm của các NST trong quá trình giảm phân. Những sai lầm này có thể phát sinh ở cả hai lần phân bào, chẳng hạn như sự không phân ly đồng đều của các NST. Sự không phân ly NST ở lần phân bào thứ nhất có thể xảy ra khi cả hai NST tương đồng cùng di chuyển về một cực ở kỳ sau I. Kết quả là trong số bốn tế bào con, hai tế bào sẽ có hai bản sao của cùng một NST, còn hai tế bào kia thì không có NST đó. Còn giả sử trong trường hợp các NST không phân ly ở kỳ sau II, thì thể tứ trị sẽ có hai tế bào chứa một NST bình thường, một tế bào thiếu một NST, còn một tế bào chứa hai bản sao của cùng một NST. Hiện tượng dị bội cũng có thể phát sinh do sự không phân ly NST ở những lần nguyên phân đầu tiên của hợp tử. Điều đó sẽ sinh ra hai quần thể tế bào có bộ NST khác nhau trong cùng một cá thể, được gọi là thể khảm. Hiện tượng này phổ biến ở hội chứng Turner.

III.D LIÊN KẾT GEN

Định nghĩa

Liên kết gen là xu hướng các alen của hai hoặc nhiều gen được di truyền theo cùng một tổ hợp từ thế hệ này sang thế hệ khác. Nhìn chung, hai gen càng gần nhau trên cùng nhiễm sắc thể (NST) thì càng có xu hướng liên kết chặt chẽ với nhau. Trên cơ sở tập tính của các NST trong quá trình phân bào, có thể dễ dàng nhận thấy các gen nằm trên các NST không tương đồng không thể liên kết với nhau, mà thay vào đó chúng phân ly độc lập ngay ở lần phân bào thứ nhất của quá trình giảm phân.

Xét một cá thể dị hợp tử kép (AaBb) có A và B là hai gen nằm trên hai NST khác nhau. Giả sử cá thể này được di truyền các alen trội (A và B) từ bố, còn các alen lặn (a và b) từ mẹ. Trong giảm phân, bốn loại giao tử khác nhau là AB, ab, Ab và aB được tạo ra với tỷ lệ tương đương. Các giao tử AB và ab được gọi là các giao tử bố, mẹ; còn các giao tử Ab và aB là các giao tử tái tổ hợp. Tần số tái tổ hợp rõ ràng là 50%. Đó là tần số tái tổ hợp cao nhất có thể có được giữa hai gen bất kỳ và các gen có tần số tái tổ hợp như vậy được gọi là các gen không liên kết.

Điều gì sẽ xảy ra nếu hai gen liên kết trên cùng một NST? Chúng sẽ chỉ được tổ hợp lại khi có trao đổi chéo. Hai gen càng gần nhau càng ít có khả năng trao đổi chéo. Vì thế, tần số tái tổ hợp sẽ là thước đo gián tiếp mức độ gần nhau của hai gen.

Trong thực tế, tần số trao đổi chéo ở hầu hết các nhiễm sắc thể là tương đối cao. Điều đó cũng có nghĩa là các gen nằm xa nhau trên cùng một NST cũng có thể có tần số tái tổ hợp là 50%. Nếu hai gen liên kết và gen thứ ba liên kết với một trong hai gen đó thì chúng phải nằm trên cùng một NST. Các gen này lập thành nhóm liên kết. Như vậy, bản đồ liên kết có thể được xây dựng cho dù một số gen trong nhóm có thể không thể hiện sự liên kết với một số gen khác trong nhóm.

Cách đo lường đơn giản sự liên kết gen

Trừ một số trường hợp có thể đo lường trực tiếp sự liên kết gen với tâm động, còn thông thường các nhà nghiên cứu cần phải có các cá thể dị hợp tử về hai gen để có thể tiến hành nghiên cứu. Lúc đó, phép lai hai tính sẽ được thực hiện như trình bày trong ví dụ sau:

Người ta tiến hành lai hai dòng cà chua thuần chủng, một dòng là đồng hợp tử quả đỏ tía và thân có lông (kiểu gen *PPHH*), còn dòng kia là đồng hợp tử quả đỏ thân nhẵn (kiểu gen *pphh*) thu được con lai có màu quả đỏ tía và thân có lông. Như vậy quả đỏ tía là tính trạng trội so với quả đỏ, và thân có lông là trội so với thân nhẵn. Để xác định xem hai gen này có liên kết hay không, con lai dị hợp tử kép được tiến hành lai với đồng hợp tử lặn kép (quả đỏ và thân không lông). Đó là phép lai phân tích. Kết quả thu được bốn lớp kiểu hình ở đời con (*Bảng 3.7*). Trong đó, lớp 1 và 2 có kiểu hình giống với bố hoặc mẹ, vì thế được gọi là các **lớp kiểu hình bố, mẹ**; lớp 3 và 4 là những tổ hợp alen mới của hai gen và vì thế được gọi là các lớp **kiểu hình tái tổ hợp**.

Nếu các gen không liên kết, tỷ lệ bốn lớp kiểu hình sẽ là 1:1:1:1. Trong thí nghiệm này, kết quả không như vậy. Tổng số cây con thu được là 500, trong đó 70 cây là các thể tái tổ hợp. Tần số tái tổ hợp là 70/500 = 14%. Đó là bằng chứng chắc chắn rằng hai gen liên kết với nhau và cách nhau 14 đơn vị bản đồ (1 đơn vị bản đồ tương ứng với 1% trao đổi chéo). Đơn vị bản đồ được gọi là centimorgan (cM) để tưởng nhớ Thomas Hunt Morgan, nhà di truyền học đầu tiên giải thích hiện tượng liên kết gen.

Một câu hỏi đặt ra là: để xác định sự liên kết gen, có nhất thiết luôn luôn tiến hành phép lai phân tích không? Câu trả lời là không nhất thiết, vì thay vào đó, người ta có thể sử dụng các phép lai khác. Ví dụ, phép lai hai dị hợp tử kép cũng cho ta các số liệu về liên kết gen vì các lớp kiểu hình tái tổ hợp từ tỷ lệ phân ly 9:3:3:1 (trong trường hợp phân ly độc lập) sẽ giảm đi. Hơn nữa, vì trong phép lai như vậy không có bố (hoặc mẹ) là đồng hợp tử lặn, nên tỷ lệ thu được chính là tỷ lệ giao tử và vì vậy phép lai này sẽ cho nhiều thông tin nhất về sự liên kết gen. Trong những nghiên cứu về liên kết gen ở người, tần số tái tổ hợp thường được xác định bằng các phương pháp phân tử nhằm phân biệt sự sai khác trong trình tự ADN (kiểu gen) chứ không nhất thiết phải dựa vào kiểu hình như ví dụ trên đây. Nghiên cứu trực tiếp các chỉ thị ADN (kiểu gen) cũng cho nhiều thông tin về phép lai. Trong thực tế, phép lai hai tính chỉ dùng để xác định sự liên kết và khoảng cách tương đối giữa hai gen. Để xác định trật tự các gen trong nhóm liên kết và tính toán chính xác khoảng cách giữa chúng trên NST cần phải sử dụng ít nhất ba gen trong một phép lai (phép lai ba tính).

Bảng 3.7. **Kiểu hình thu được trong phép lai phân tích cây cà chua F1 quả đỏ tía, thân có lông (*Pp, Hh*) với cây bố/mẹ đồng hợp tử về các alen lặn quả đỏ, thân không lông (*pp, hh*)**

Thứ tự lớp kiểu hình	Các lớp kiểu hình	Số lượng
(1)	Quả đỏ tía, thân có lông	220
(2)	Quả đỏ, thân không lông	210
(3)	Quả đỏ tía, thân không lông	32
(4)	Quả đỏ, thân có lông	38
	Tổng số	500

Phép lai ba tính

Ưu điểm chính của phép lai ba tính là giúp xác định trình tự gen và khoảng cách giữa các gen có thể cộng kế tiếp. Điều đó được thể hiện trong phép lai phân tích ở ví dụ dưới đây.

Người ta tiến hành lai một cá thể dị hợp tử về ba gen *Aa, Nn, Rn* với cá thể đồng hợp tử lặn về ba gen *aa, nn, rr*. Tần số các kiểu hình ở đời con thu được như trình bày ở *Bảng 3.8*.

Bảng 3.8. **Tần số các kiểu hình đời con**

Kiểu hình	Tần số	Thứ tự lớp kiểu hình
Bố, mẹ		
A N R	347	(1)
a n r	357	(2)
Tái tổ hợp		
A N r	52	(3)
a n R	49	(4)
A n r	90	(5)
a N R	92	(6)
A n R	6	(7)
a N r	7	(8)
Tổng số	1000	

Nếu các gen không liên kết thì tám nhóm kiểu hình phải có tần số như nhau. Kết quả ở đây không như vậy: lớp 1 và 2 có kiểu hình giống bố, mẹ, vậy chúng không phải là đời con do các giao tử tái tổ hợp tạo nên; lớp 3 và 4 là các kiểu hình do các giao tử tái tổ hợp giữa gen *N* và *R* tạo nên; lớp 5 và 6 là do tái tổ hợp giữa gen *A* và *N*; lớp 7 và 8 rất quan trọng vì chúng là kết quả được tạo ra do tái tổ hợp giữa hai cặp gen *A* và *N*, và *N* và *R*. Hiện tượng tạo ra lớp kiểu hình 7 và 8 này được gọi là trao đổi chéo kép, xảy ra với tần số rất thấp vì nó cần hai sự kiện trao đổi chéo (TĐC) xảy ra độc lập, đồng thời. Vì vậy, hai lớp kiểu hình có tần số thấp nhất ở đời con sẽ được dùng để xác định gen nào nằm giữa trong ba gen liên kết. Trong trường hợp này, gen nằm giữa là gen *N*.

Để xác định khoảng cách trên bản đồ giữa hai gen, cần phải tính tần số tái tổ hợp giữa các cặp gen, cụ thể gồm các bước sau:

1) Xác định khoảng cách giữa gen A và N:

– Cộng số cá thể lớp kiểu hình 5 và 6, là kết quả do tái tổ hợp giữa *A* và *N*.

– Cộng số cá thể lớp kiểu hình 7 và 8, là kết quả do trao đổi chéo kép, trong đó một trao đổi chéo là giữa *A* và *N*.

– Biểu diễn tổng số con do trao đổi chéo sinh ra bằng %:

$$\frac{(90+92+6+7)\times 100}{1000} = 19,5\%, \text{ hay } 19,5\text{cM}$$

2) Xác định khoảng cách giữa gen N và R:

- Tương tự như trên, khoảng cách giữa gen N và R bằng:

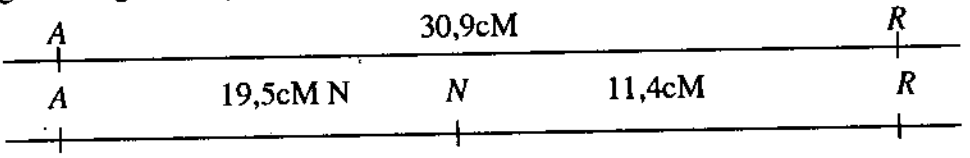
$$\frac{(52+49+6+7)\times 100}{1000} = 11,4\%, \text{ hay } 11,4\text{cM}$$

3) Xác định khoảng cách giữa gen A và R:

– Tính tần số tất cả các thể tái tổ hợp:

$$\frac{(52+49+90+92+6+7+6+7)\times 100}{1000} = 30,9\%, \text{ hay } 30,9\text{cM}$$

Trên cơ sở khoảng cách giữa các gen được cộng kế tiếp, ta suy ra trật tự và ước tính được khoảng cách giữa các gen trong nhóm liên kết này là:



Phép lai ba tính chặt chẽ hơn phép lai hai tính vì nó giúp phát hiện và sử dụng tần số TĐC kép. Tuy vậy, các nghiên cứu di truyền gần đây cho thấy: tần số tái tổ hợp giữa các gen trong nhóm liên kết qua kết quả phân tích các lớp kiểu hình trong các phép lai không luôn luôn tỷ lệ tuyến tính với khoảng cách giữa các gen trên NST, hay nói cách khác, tần số tái tổ hợp trong trường hợp này chỉ đơn thuần phản ánh tần số TĐC chứ không phải là thước đo khoảng cách vật lý thực sự giữa các gen. Quan điểm này sẽ còn được đề cập ở mục VI.E.

Nhiễu

Trao đổi chéo xảy ra tại các cấu trúc vắt chéo giữa các nhiễm sắc tử trong cặp NST tương đương. Nhưng có một điểm đáng chú ý là: sự có mặt của một cấu trúc vắt chéo tại một vùng nào đó trên NST có thể ảnh hưởng và làm giảm tần số xuất hiện các cấu trúc vắt chéo khác ở gần nó. Vì vậy, tần số TĐC kép thực tế thường thấp hơn tần số TĐC kép lý thuyết. Ở ví dụ trên, chúng ta đã thấy tần số TĐC kép là 13/1000 (hay 0,013). Bằng cách sử dụng các tần số TĐC đơn giữa *A* và *N*, giữa *N* và *R*, chúng ta có thể tính được tần số TĐC kép theo lý thuyết như sau:

– Số trao đổi chéo giữa *A* và *N* = 19,5

– Số trao đổi chéo đơn giữa *N* và *R* = 11,4

– Tần số trao đổi chéo kép theo lý thuyết (trên 1000 cá thể) sẽ là:

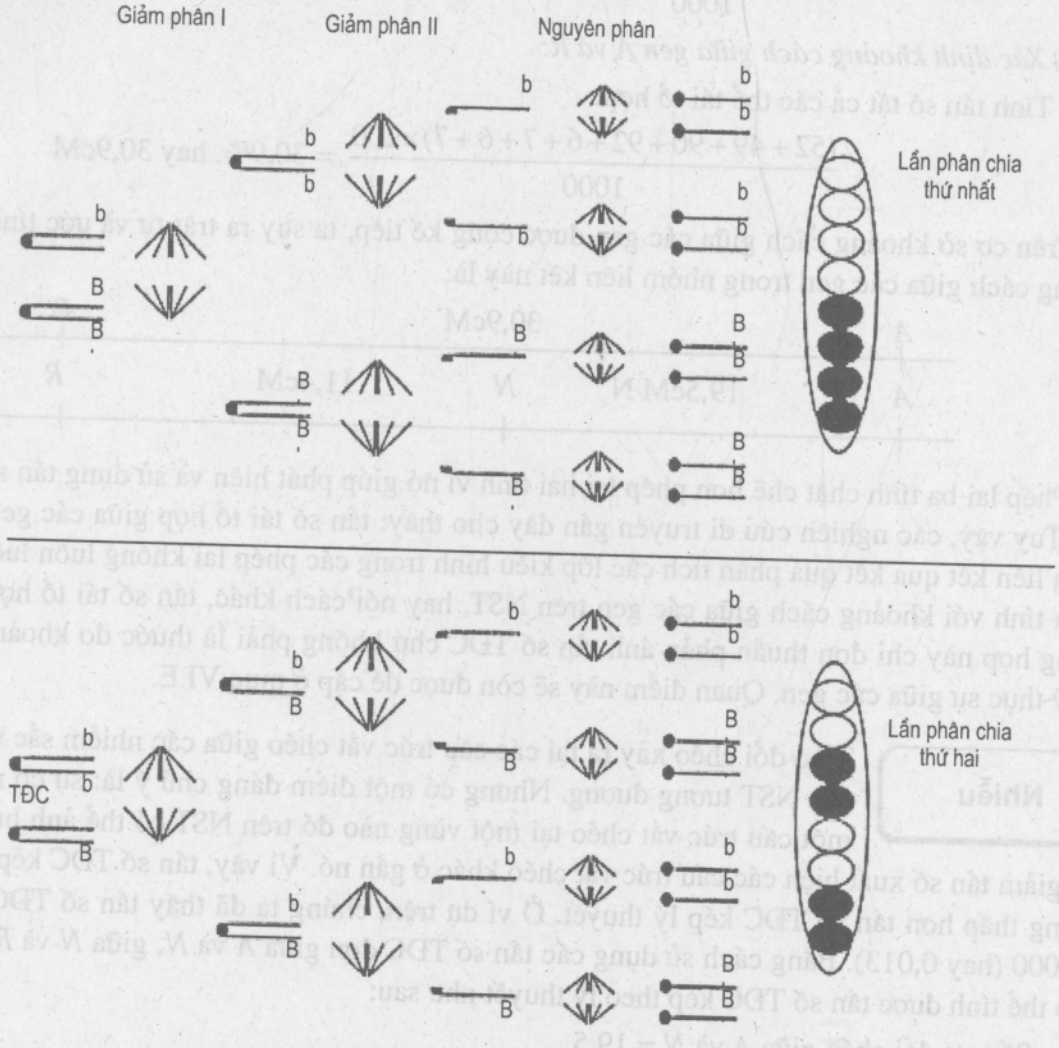
$$(195/1000) \times (114/1000) \times 1000 = 22,2$$

Như vậy, tần số TĐC kép lý thuyết cao hơn nhiều so với con số mà ta quan sát được là 13. Điều đó cho thấy có **hiện tượng nhiễu dương tính**. Mức độ nhiễu được biểu diễn bằng **hệ số nhiễu S**. Hệ số này được tính bằng tỷ số giữa số TĐC kép quan sát được chia cho số TĐC kép tính theo lý thuyết. Trong trường hợp nêu trên, hệ số S sẽ là:

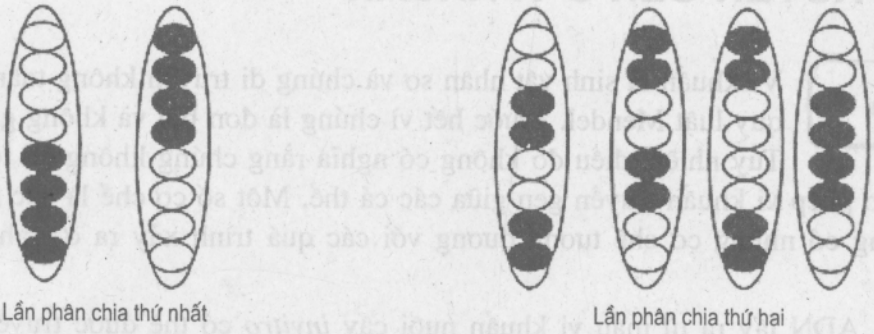
$$S = 13/22,2 = 0,58$$

Phân tích liên kết gen ở nấm

Một số nấm nang sinh ra các nang chứa các bào tử đơn bội do giảm phân sinh ra và được xếp theo một trật tự đặc biệt gọi là bộ bốn xếp hàng. Trật tự của bộ bốn này phản ánh đúng trật tự của các thể lưỡng trị trong giảm phân. Mỗi cặp bào tử là sản phẩm của một thể lưỡng trị sau hai lần phân chia giảm phân. Ở một số loài, nang chứa tám bào tử, là vì mỗi bào tử lại trải qua một lần nguyên phân nữa. *Hình 3.12* minh họa các sản phẩm giảm phân của một nấm dị hợp tử về tính trạng bào tử màu đậm và màu nhạt, mỗi nang chứa 8 bào tử.



Hình 3.12. Giảm phân ở nấm nang dị hợp tử về gen quy định màu sắc bào tử. Alen B và b tương ứng quy định bào tử sẫm màu và nhạt màu. Phần trên là kiểu phân ly khi không có TĐC, đó là kiểu phân ly của lần phân chia thứ nhất. Phần dưới là kiểu phân ly khi có một TĐC đơn xảy ra giữa gen và tâm động, đó là kiểu phân ly của lần phân chia thứ hai



Hình 3.13. **Các kiểu sắp xếp bào tử trong nang sau lần phân chia thứ nhất và thứ hai. Sau lần phân chia thứ nhất có hai kiểu sắp xếp tùy thuộc vào vị trí của các thành viên thể lưỡng trị trong giảm phân. Kiểu sắp xếp các bào tử sau lần phân chia thứ hai phụ thuộc vào cặp nhiễm sắc tử nào tham gia vào TĐC**

Những nang có bốn bào tử cùng màu nằm liền nhau là do không có TĐC giữa gen quy định màu sắc bào tử và tâm động. Kiểu phân bố như vậy được gọi là kiểu phân ly của lần phân chia thứ nhất vì đó là sự phân tách vật lý ở giảm phân I. Để có các kiểu phân ly khác phải có TĐC xảy ra giữa tâm động và gen quy định màu sắc bào tử. Khi đó kiểu sắp xếp màu sắc của bào tử trong nang sẽ khác đi sau giảm phân II, đó được gọi là kiểu phân ly của lần phân chia thứ hai. Có thể có các kiểu sắp xếp khác nhau của các bào tử trong nang phụ thuộc vào nhiễm sắc tử nào tham gia trao đổi chéo. Các kiểu sắp xếp đó được trình bày trên *Hình 3.13*.

Tần số tái tổ hợp được tính theo công thức như sau:

$$\frac{\text{(1/2 số nang sắp xếp theo kiểu phân ly của lần phân chia thứ hai)} \times 100}{\text{(Tổng số nang được sinh ra)}}$$

Chỉ lấy1/2 số nang sắp xếp theo kiểu phân ly của lần phân chia thứ hai vì chỉ có một nửa số nhiễm sắc tử của các nang đó tham gia TĐC giữa gen quy định màu sắc bào tử và tâm động. Khoảng cách giữa hai gen cũng có thể được xác định theo cách tương tự. Ngoài ra, nếu nhà nghiên cứu không muốn tách nang thì có thể xác định khoảng cách trên bản đồ bằng cách tính số nang không tạo bộ bốn xếp hàng dựa theo kiểu hình của các bào tử thu được từ mỗi nang. Chỉ có cách phân tích bộ bốn xếp hàng mới xác định được khoảng cách giữa gen và tâm động.

Tần số tái tổ hợp và khoảng cách vật lý

Trước đây, tần số tái tổ hợp thường được coi là thước đo khoảng cách thực sự giữa các gen. Điều này không phải luôn luôn đúng. Hiện nay, chúng ta đã biết rằng tần số TĐC giữa các gen không hoàn toàn tỷ lệ thuận với khoảng cách giữa chúng trên NST. Có những vùng trên NST có tần số TĐC cao hơn so với các vùng khác. Chẳng hạn, tần số TĐC thường thấp hơn ở các vùng gần tâm động và cao hơn ở vùng gần đầu mút của NST. Tần số TĐC còn khác nhau giữa giới đực và giới cái. Tuy vậy, bản đồ gen được xác định dựa trên các số liệu phân tích tái tổ hợp kiểu hình như nêu ở trên vẫn cung cấp cho chúng ta những chỉ dẫn khá chính xác về khoảng cách giữa các gen, và là phương pháp được sử dụng rộng rãi và hiệu quả nhất cho đến nay để xác định trật tự của các gen trên nhiễm sắc thể.

III.E SỰ TRUYỀN GEN Ở VI KHUẨN

Tổng quan

Vi khuẩn là sinh vật nhân sơ và chúng di truyền không tuân theo các quy luật Mendel. Trước hết vì chúng là đơn bội và không giảm phân. Tuy nhiên, điều đó không có nghĩa rằng chúng không tái tổ hợp. Có nhiều cơ chế cho phép vi khuẩn truyền gen giữa các cá thể. Một số cơ chế là đặc trưng ở vi khuẩn, song cũng có những cơ chế tương đương với các quá trình xảy ra ở sinh vật nhân chuẩn.

Ở vi khuẩn, ADN lấy ra từ mẫu vi khuẩn nuôi cấy *invitro* có thể được truyền sang vi khuẩn thể nhận. Quá trình đó được gọi là sự **biến nạp**, hay sự truyền nhiễm. Vi khuẩn có thể được cải tiến để làm tăng tỷ lệ các tế bào **khả biến** có thể tiếp nhận ADN ngoại lai. Sau đó, chúng được trộn với dịch ADN ngoại lai tinh sạch, và những tế bào biến nạp (đã nhận được ADN) được chọn lọc trên các môi trường nuôi cấy chọn lọc. Hiện tượng biến nạp rất có ích để chọn tạo ra được các chủng vi khuẩn có được những đặc điểm di truyền đặc biệt, hoặc nhằm mục đích chuyển gen giữa các loài. Biến nạp còn có thể được dùng để xác định và lập bản đồ liên kết gen. Điều này được thực hiện bằng cách so sánh tần số biến nạp của từng gen với tần số đồng biến nạp của chúng.

Chẳng hạn, nếu gen *A* được truyền sang vi khuẩn thể nhận với tần số 10^{-4} và gen *B* với tần số 4×10^{-5} thì tần số đồng biến nạp của hai gen trên lý thuyết sẽ là 4×10^{-9} (bằng tích của tần số biến nạp của hai gen). Đó là dựa theo giả thiết rằng mỗi sự kiện biến nạp là sự lấy vào một đoạn ADN, hoặc mang gen *A*, hoặc mang gen *B*. Tuy nhiên, nếu hai gen *A* và *B* nằm rất gần nhau trên nhiễm sắc thể của vi khuẩn thể cho thì tỷ lệ đoạn ADN mang cả hai gen đồng biến nạp tăng lên. Do vậy, tần số đồng biến nạp cũng sẽ tăng. Quá trình này cho phép phân biệt các gen liên kết gần. Đối với những gen ở quá xa nhau, có thể không xảy ra hiện tượng đồng biến nạp, nhưng vì mỗi gen trong số đó có thể liên kết với gen khác nữa nên cả ba gen đều phải nằm gần nhau trên nhiễm sắc thể của thể cho. Như vậy, việc lập bản đồ gen có thể được xây dựng theo cách giống với phương pháp lập bản đồ di truyền dựa trên tần số trao đổi chéo ở sinh vật nhân chuẩn.

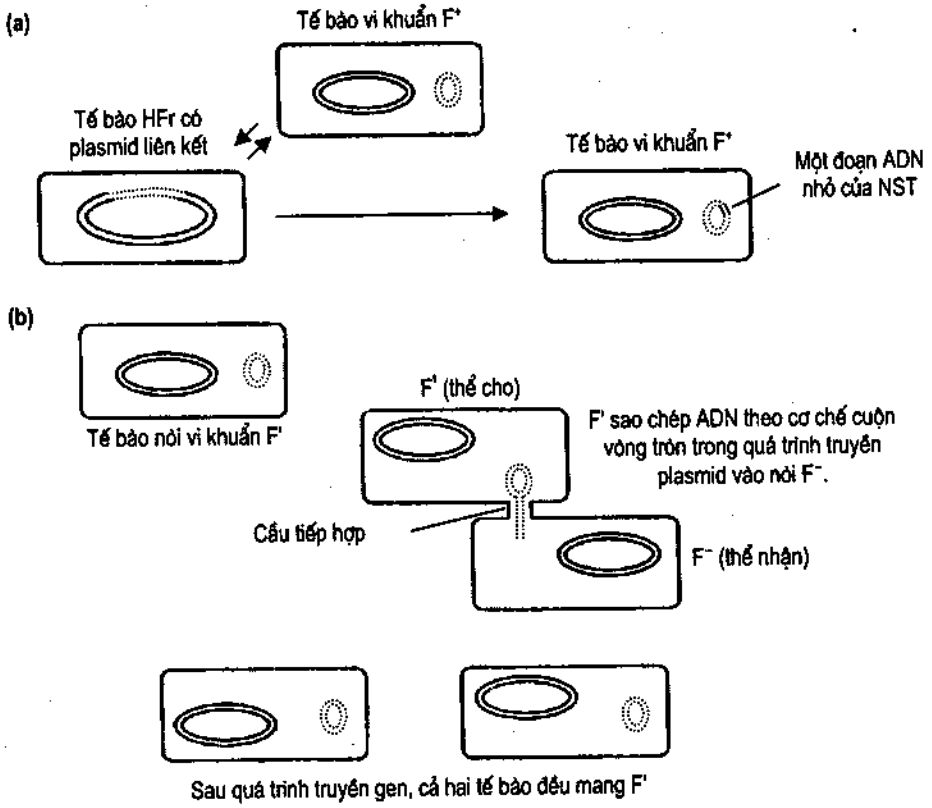
Hiện tượng biến nạp, cho dù có thể xảy ra trong tự nhiên, nhưng hiện nay được ứng dụng nhiều trong các nghiên cứu ở phòng thí nghiệm, chẳng hạn để lập bản đồ gen ở các loài vi khuẩn. Ngoài hiện tượng biến nạp, ở vi khuẩn còn có các quá trình truyền gen tự nhiên có thể được dùng để lập bản đồ gen. Các quá trình đó, hoặc phụ thuộc vào hoạt động của **thực khuẩn thể (bacteriophage/phage)** gọi là **quá trình tải nạp**, hoặc liên quan đến hoạt động của một cấu trúc giống như nhiễm sắc thể là **plasmid,** gọi là **quá trình tiếp hợp.**

Plasmid là trung gian truyền gen

Plasmid có ở nhiều loài vi khuẩn. Chúng được cấu tạo từ chuỗi ADN mạch kép mã hoá protein giống với nhiễm sắc thể vi khuẩn. Tuy nhiên, plasmid có thể bị tách khỏi vi khuẩn mà không gây hại gì cho tế bào chủ. Vì vậy, rõ ràng chúng là những cấu trúc không thiết yếu. Trong nhiều trường hợp, plasmid có thể vận động từ tế bào vi khuẩn này sang tế bào vi khuẩn khác, thậm chí giữa các loài vi khuẩn khác nhau. Quá trình truyền gen được điều khiển bởi các gen trên plasmid. Khả năng vận động giữa các vi khuẩn của plasmid có thể được dùng để nghiên cứu sự truyền gen của nhiễm sắc thể ở vi khuẩn.

Plasmid đầu tiên được chứng minh có khả năng vận động giữa các vi khuẩn là plasmid F. Plasmid này mang các gen điều khiển sự hình thành lông F (còn gọi là lông giới tính, hay ống tiếp hợp). Cấu trúc này được tạo ra giữa các vi khuẩn mang plasmid (gọi là nòi F^+) với các vi khuẩn không có plasmid (F^-) để plasmid F có thể truyền qua. Plasmid này nhân đôi bằng cách sao chép cuộn vòng tròn khi nó được truyền đi. Điều đó có nghĩa là, tế bào thể cho giữ lại một bản sao plasmid và khi trộn vi khuẩn F^+ với F^- thì plasmid F sẽ lan truyền cho mọi tế bào trong đĩa nuôi cấy. Đôi khi plasmid được gọi là yếu tố F hoặc yếu tố giới tính và vi khuẩn F^+ được coi là giới đực.

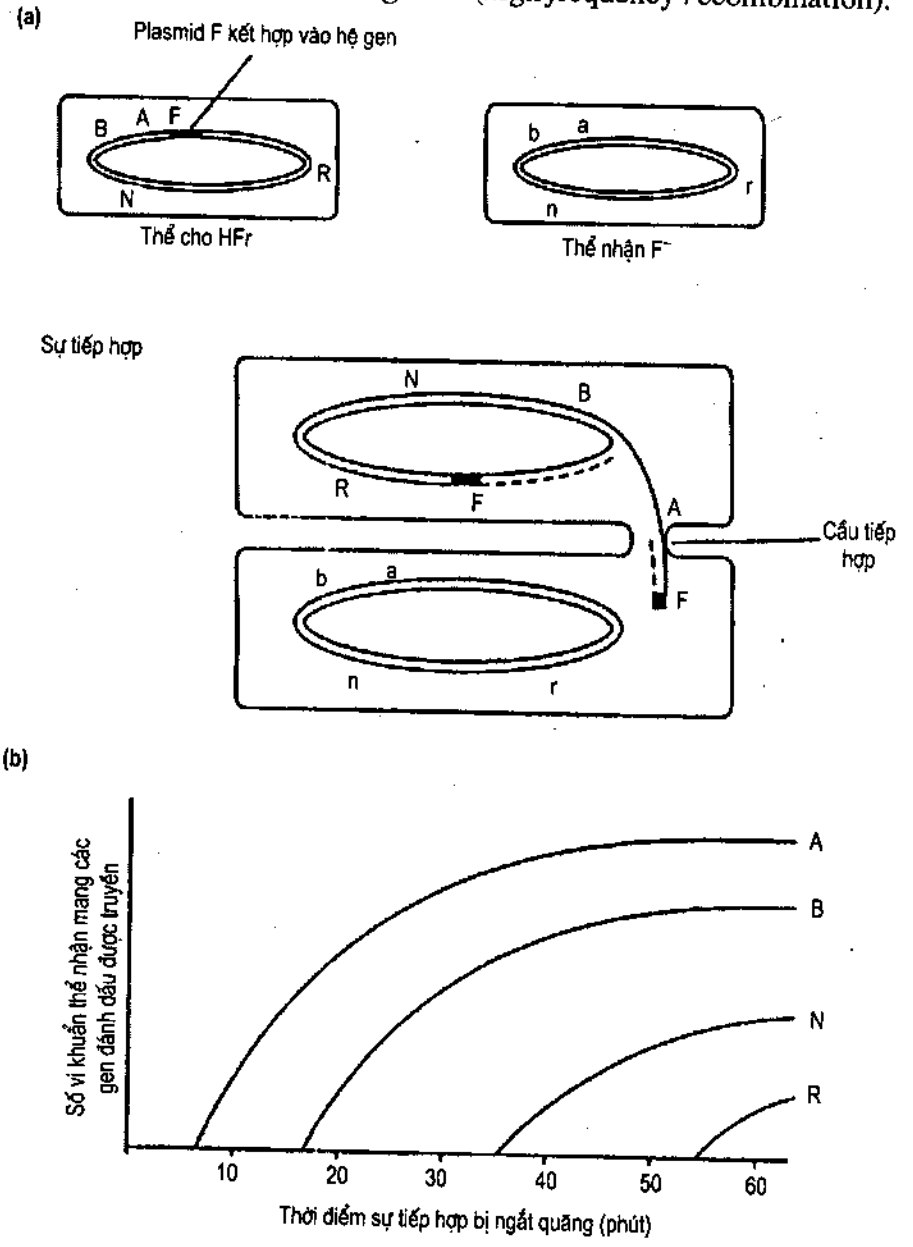
Plasmid F có thể tồn tại độc lập. Nó sao chép độc lập với NST vi khuẩn. Nó cũng có thể lồng ghép vào nhiễm sắc thể vi khuẩn và sao chép như một phần của nhiễm sắc thể đó. Chúng ta gọi các plasmid có thể tồn tại được ở hai trạng thái như vậy là các episom. Quá trình này là thuận nghịch và đôi khi plasmid F bị cắt khỏi nhiễm sắc thể. Đôi khi quá trình cắt bỏ này bị nhầm, sinh ra các plasmid mang các đoạn nhiễm sắc thể vật chủ có độ dài ngắn khác nhau, khi đó nó được gọi là F' hay yếu tố giới tính thứ cấp. Quá trình đó được thể hiện trên *Hình 3.14a*. Yếu tố F' có khả năng vận động sang vi khuẩn F^-, mang theo cả một số gen của vi khuẩn chủ. Hiện tượng đó được gọi là sự truyền gen giới tính (*Hình 3.14b*). Sự truyền gen nhờ quá trình này xảy ra với tần số thấp.



Hình 3.14. **(a) Mối liên quan giữa plasmid F và nhiễm sắc thể vi khuẩn. ADN của plasmid là đường cong gạch nối, ADN nhiễm sắc thể là đường cong liền. (b) Quá trình truyền plasmid F' từ vi khuẩn F' (thể cho) sang vi khuẩn F^- (thể nhận)**

Tần số truyền gen cao hơn nhiều khi plasmid F được lồng ghép vào trong NST. Nó không những vẫn duy trì khả năng truyền yếu tố F sang F^- mà còn mang theo phần còn lại của NST vi khuẩn. Quá trình này được gọi là **quá trình tiếp hợp** (*Hình 3.15a*), diễn ra với một số bước

như sau: NST vi khuẩn bị đứt gãy và được truyền sang vi khuẩn thể nhận F^- cùng với yếu tố F; vị trí yếu tố F được truyền sang đầu tiên được gọi là điểm khởi đầu truyền gen; phần còn lại của yếu tố F được truyền sang sau cùng. Cần khoảng gần 2 giờ để truyền toàn bộ nhiễm sắc thể vi khuẩn sang thể nhận, nhưng hiện tượng này ít khi đạt được vì ống tiếp hợp dễ bị gãy. Hiện tượng tiếp hợp này tạo ra sự truyền gen ở vi khuẩn với tần số cao. Vì thế vi khuẩn mang yếu tố F lồng ghép được gọi là các **chủng HFr** (*h*igh *f*requency *r*ecombination).



Hình 3.15. (a) Vi khuẩn thể cho HFr và vi khuẩn thể nhận F^-. HFr mang plasmid F được lồng ghép vào NST của nó. Bốn gen đánh dấu A, B, N và R để chọn lọc được thể hiện ở các vị trí tương ứng trên bản đồ gen. Thể nhận mang các gen lặn của bốn gen đó (ký hiệu bằng chữ thường). Trong quá trình tiếp hợp, NST bị gãy bên trong yếu tố F. Sợi đơn ADN được truyền qua ống tiếp hợp. Chú ý rằng, sợi này được sao chép bên trong thể nhận và quá trình tổng hợp sợi mới thứ hai cũng được bắt đầu ở thể cho ở vùng cách xa yếu tố F, được biểu diễn bằng đường chấm. Đó là sự sao chép vòng lăn. (b) Trật tự và vị trí của các gen A, B, N và R được xác định bằng cách lập đồ thị số lượng vi khuẩn nhận mang các dấu chuẩn chọn lọc tại các thời điểm khác nhau khi sự tiếp hợp bị ngắt quãng. Đồ thị cho thấy vị trí tương đối của các gen đánh dấu là: A – 8 phút; B – 17 phút; N – 36 phút; R – 54 phút kể từ điểm khởi đầu truyền gen.

Vì sự truyền gen thường được khởi đầu từ một điểm cố định, và vì NST trong quá trình tiếp hợp được truyền sang thể nhận dưới dạng một cấu trúc mạch thẳng với tốc độ tương đối ổn định nên trật tự và thời điểm các gen cụ thể xâm nhập vào thể nhận F^- có thể được dùng để lập bản đồ NST ở vi khuẩn. Sau khi trộn các vi khuẩn HFr và F^- và cho chúng tiếp hợp thì có thể dừng tiếp hợp ở những thời điểm khác nhau bằng cách gây rung mạnh môi trường nuôi cấy để các ống tiếp hợp đứt gãy, ngừng quá trình truyền NST. Quá trình này được gọi là **giao phối ngắt quãng**. Thời điểm một gen được truyền sang có thể xác định được nếu lấy vi khuẩn từ môi trường nuôi cấy chuyển sang môi trường chọn lọc sau những khoảng thời gian khác nhau trong quá trình thí nghiệm. Như vậy, bản đồ gen của vi khuẩn được xây dựng thành bản đồ thời gian (tính bằng phút), cho thấy khoảng cách mỗi gen được truyền sang kể từ điểm khởi đầu truyền gen. Ví dụ minh hoạ phương pháp này được minh họa trên *Hình 3.15b*.

Thực khuẩn thể (phage) và sự truyền gen

Thực khuẩn thể cũng có thể được dùng trong quá trình truyền gen ở vi khuẩn. Các phage độc sao chép trong tế bào vi khuẩn. Một hạt phage thông thường ngoài ADN của chính nó có thể mang một đoạn nhỏ ADN của vi khuẩn. Phage lúc đó sẽ có khả năng truyền đoạn gen này vào tế bào thể nhận.

Phage ôn hòa có khả năng liên kết hệ gen của nó vào NST vi khuẩn và duy trì ở trạng thái tiềm tan. Sự kết hợp hệ gen chỉ xảy ra ở một vị trí đặc hiệu. Trạng thái tiềm tan mất đi ở một tần số thấp, khi đó hệ gen của phage được tách khỏi hệ gen tế bào chủ và bắt đầu sao chép. Trong một số trường hợp, hiện tượng hệ gen của phage tách khỏi tế bào vi khuẩn không chính xác và dẫn đến sự hình thành các phage mang theo một phần phân tử ADN của vi khuẩn. Đoạn phân tử ADN này sau đó có thể được truyền sang một tế bào chủ khác khi phage tiếp tục lây nhiễm. Thực khuẩn thể chỉ có thể kết hợp với phân tử ADN tế bào chủ ở những vị trí tiếp giáp với vị trí kết hợp đặc hiệu của nó, vì vậy quá trình này gọi là quá trình truyền gen hạn chế.

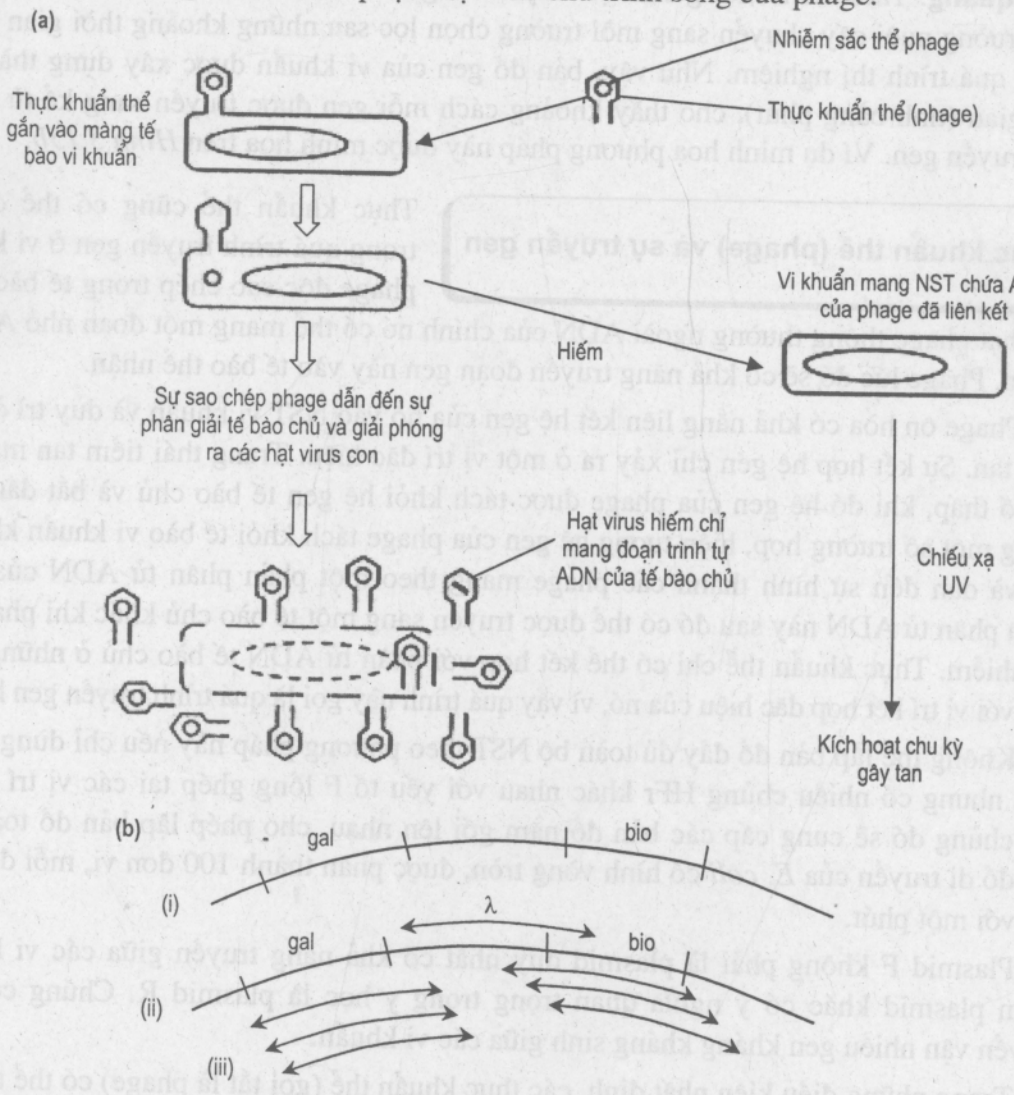
Không thể lập bản đồ đầy đủ toàn bộ NST theo phương pháp này nếu chỉ dùng một chủng HFr, nhưng có nhiều chủng HFr khác nhau với yếu tố F lồng ghép tại các vị trí khác nhau. Các chủng đó sẽ cung cấp các bản đồ nằm gối lên nhau, cho phép lập bản đồ toàn bộ NST. Bản đồ di truyền của *E. coli* có hình vòng tròn, được phân thành 100 đơn vị, mỗi đơn vị tương ứng với một phút.

Plasmid F không phải là plasmid duy nhất có khả năng truyền giữa các vi khuẩn. Một nhóm plasmid khác có ý nghĩa quan trọng trong y học là plasmid R. Chúng có khả năng chuyển vận nhiều gen kháng kháng sinh giữa các vi khuẩn.

Trong những điều kiện nhất định, các thực khuẩn thể (gọi tắt là phage) có thể thực hiện sự truyền gen giữa các vi khuẩn. Quá trình đó được gọi là **tải nạp**. Quá trình này liên quan đến những sai sót trong quá trình sao chép của phage và rất hữu ích để lập bản đồ các gen ở vi khuẩn.

Các phage có thể được phân thành hai nhóm là phage độc và phage ôn hoà, phụ thuộc vào hoạt động của chúng sau khi nhiễm vào tế bào vi khuẩn. Khi một phage độc nhiễm vào tế bào vi khuẩn, nó sử dụng bộ máy tổng hợp của tế bào chủ để sản sinh các hạt phage mới. Tiếp đó, vi khuẩn bị phân giải, giải phóng ra một lượng lớn phage (chu trình này gọi là chu trình gây tan). Các phage ôn hoà phải lựa chọn giữa chu trình gây tan và chu trình tiềm tan. Con đường tiềm tan bao gồm sự lồng ghép hệ gen của phage vào nhiễm sắc thể vật chủ, khi đó nó được

gọi là tiền phage, và sao chép cùng với NST vật chủ. Phage λ là một ví dụ về phage ôn hoà. Phage này chỉ có thể lồng ghép vào NST vật chủ tại một vị trí đặc thù nằm giữa gen *gal* và gen *bio* trên NST vi khuẩn. Khi đó, vi khuẩn không bị ảnh hưởng gì nhưng cũng không bị nhiễm tiếp các phage cùng loại. Tiền phage có thể mất trạng thái ngủ trên đây và chuyển sang chu trình gây tan (với tần số thấp), tiến hành sao chép vật chất di truyền và làm tan tế bào vi khuẩn. Vi khuẩn tiềm tan có thể bị cảm ứng thành gây tan bởi các tác nhân như tia cực tím. Phương thức truyền các gen vi khuẩn phụ thuộc vào chu trình sống của phage.



(i) Vị trí các trình tự gen *gal* và *bio* trên một phần NST của vi khuẩn *E. coli*.
(ii) Sự xen trình tự ADN của phage λ vào giữa các gen của vi khuẩn.
(iii) Thành phần ADN của các hạt virus mang các gen không hoàn chỉnh do chiếu xạ UV gây ra, hoặc chỉ mang một phần các trình tự gen *gal* và *bio*, và một phần hệ gen λ.

Hình 3.16. **(a) Chu trình gây tan sinh ra các hạt tải nạp của hiện tượng tải nạp chung. (b) Hiện tượng tải nạp đặc hiệu vùng gen *gal - bio* ở *E. coli***

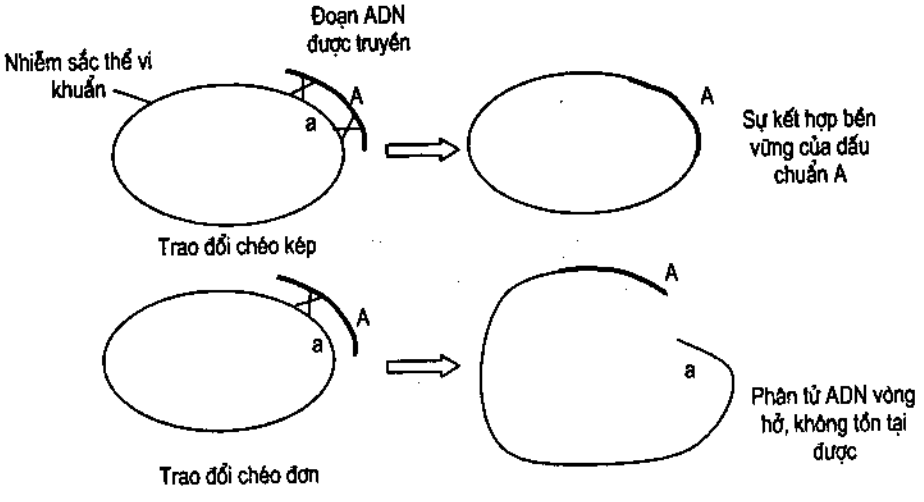
Trong chu trình gây tan của một số phage độc, sai sót có thể xảy ra, dẫn đến một đoạn nhỏ NST vi khuẩn chủ được lồng ghép ngẫu nhiên vào ADN của các hạt phage mới. Cấu trúc như vậy được gọi là các hạt tải nạp. *Hình 3.16a* mô tả sự tương tác giữa phage và vi khuẩn để

sản sinh ra các hạt tải nạp. Hạt tải nạp chỉ xuất hiện với tần số thấp (khoảng 10^{-6}). Chúng vẫn có khả năng gây nhiễm vi khuẩn vì có chứa các protein vỏ phage, nhưng vì chúng mang một phần ADN vi khuẩn nên chúng đồng thời có thể hoạt động như một vector truyền ADN giữa các vi khuẩn. Các môi trường chọn lọc thích hợp sẽ chỉ cho phép những vi khuẩn mang các gen tái tổ hợp theo cách này sinh trưởng, qua đó có thể phát hiện được hiện tượng tải nạp. Do có kích thước nhỏ nên phage chỉ mang được đoạn ADN nhỏ và chỉ những gen nằm rất gần nhau mới có thể **đồng tải nạp** được theo cách này. Tần số đồng tải nạp có thể được dùng để lập bản đồ chi tiết các gen vi khuẩn. Vì bất kỳ đoạn nào của NST vi khuẩn cũng có thể được tải nạp theo cách này nên quá trình đó được gọi là **tải nạp chung**. Ví dụ về các phage được dùng cho tải nạp chung là P22 ở *Salmonella* và P1 ở *E. coli*.

Khi phage ôn hoà mất trạng thái ngủ và bắt đầu sao chép (chuyển từ chu trình tiềm tan sang chu trình gây tan), nó bị cắt khỏi ADN vi khuẩn chủ. Thông thường, quá trình cắt bỏ là chính xác. Tuy nhiên, đôi khi quá trình này không chính xác dẫn đến một đoạn nhỏ ADN vi khuẩn bị cắt theo và một đoạn ADN của phage tương ứng ở lại trên NST vật chủ. Phage được tạo ra theo cách đó mất đi một số gen và không thể hoàn thành toàn bộ chu trình sống, nhưng giống như các hạt tải nạp, chúng có thể truyền gen giữa các vi khuẩn. Vì các phage ôn hoà chỉ có thể lồng ghép vào vi khuẩn ở một hoặc vài điểm, nên rõ ràng chúng chỉ có thể truyền các gen của những vùng đó. Quá trình đó được gọi là tải nạp đặc hiệu. Ví dụ, phage λ có thể được dùng để lập bản đồ chi tiết vùng gen *gal - bio*. Sự tải nạp đặc hiệu như vậy được minh họa trên *Hình 3.16b*.

Tái tổ hợp ở vi khuẩn

Các cơ chế mô tả trên đây đều có liên quan đến việc truyền các đoạn NST giữa các vi khuẩn. Vi khuẩn thể nhận có thể trở thành lưỡng bội một phần và được gọi là hợp tử từng phần. Sự tái tổ hợp (việc tạo ra một kiểu gen mới, ổn định) phụ thuộc vào trao đổi chéo giữa NST vật chủ và đoạn ADN ngoại lai được truyền. Một điểm quan trọng cần lưu ý đối với quá trình này là trong khi các trao đổi chéo đơn ở sinh vật nhân chuẩn rất hữu ích để lập bản đồ gen thì ở các hợp tử từng phần luôn phải có hơn một trao đổi chéo. Đó là vì NST vi khuẩn có dạng vòng tròn. Trao đổi chéo đơn sẽ chỉ sinh ra NST mạch thẳng, làm cho vi khuẩn không sống sót được. Điều đó được minh hoạ trên *Hình 3.17*.



Hình 3.17. **Sự cần thiết phải có hơn một trao đổi chéo để lồng ghép NST trong hợp tử từng phần**

III.F SỰ DI TRUYỀN NGOÀI NHÂN

Hệ gen ty thể và lạp thể

Các tế bào nhân chuẩn có hơn một hệ gen. Hầu hết các nghiên cứu Di truyền học tập trung nghiên cứu hệ gen nhân, song ty thể và lạp thể cũng chứa ADN riêng của chúng. Hầu hết các dẫn liệu đến nay cho thấy các bào quan này nhiều khả năng có nguồn gốc từ các sinh vật nhân sơ bị các sinh vật nhân chuẩn nguyên thuỷ "bắt giữ" và từ đó chúng tồn tại trong mối quan hệ cộng sinh. Hệ gen bào quan khác hệ gen nhân chủ yếu ở chỗ: chúng giống tổ tiên nhân sơ của chúng. Hệ gen ty thể và lạp thể của một số loài đã được phân tích và giải trình tự hoàn toàn. Hầu hết ADN lạp thể của các loài thực vật có dạng mạch vòng với kích thước khoảng 50 - 200kb. Trong một lạp thể có thể có nhiều bản sao ADN hệ gen của nó. Hệ gen ty thể cũng có cấu trúc dạng mạch vòng nhưng biến động lớn hơn về kích thước giữa các loài. ADN ty thể người dài khoảng 16,6kb, trong khi của nấm men có chiều dài lớn hơn khoảng 5 lần. Thực vật có hệ gen ty thể khoảng 1Mb.

Cả ty thể và lạp thể đều có các gen mã hóa rARN và tARN riêng. Ribosom của chúng có kích thước giống với ribosom của sinh vật nhân sơ. Một số protein tìm thấy ở các bào quan này được tổng hợp từ mARN được phiên mã từ ADN của bào quan với bộ máy dịch mã (ribosom) của riêng chúng. Tuy vậy, hầu hết các protein khác của ty thể do gen nhân mã hoá và được vận chuyển vào ty thể từ tế bào chất. Sự phụ thuộc của ty thể vào sự biểu hiện của hệ gen nhân cho thấy dường như có sự đồng tiến hoá của các bào quan với tế bào nhân chuẩn nguyên thủy.

Sự di truyền theo dòng mẹ

Dù có một số ngoại lệ, song hầu hết hạt phấn của các loài thực vật không truyền ty thể và lạp thể cho hợp tử. Ở động vật, tinh trùng cũng không đóng góp đáng kể ty thể cho hợp tử. Vì lý do đó mà các kiểu hình do hệ gen bào quan mã hoá chủ yếu được di truyền từ mẹ. Hiện tượng này gọi là **sự di truyền theo dòng mẹ**. Có rất nhiều ví dụ về hiện tượng này, chẳng hạn như các ví dụ dưới đây:

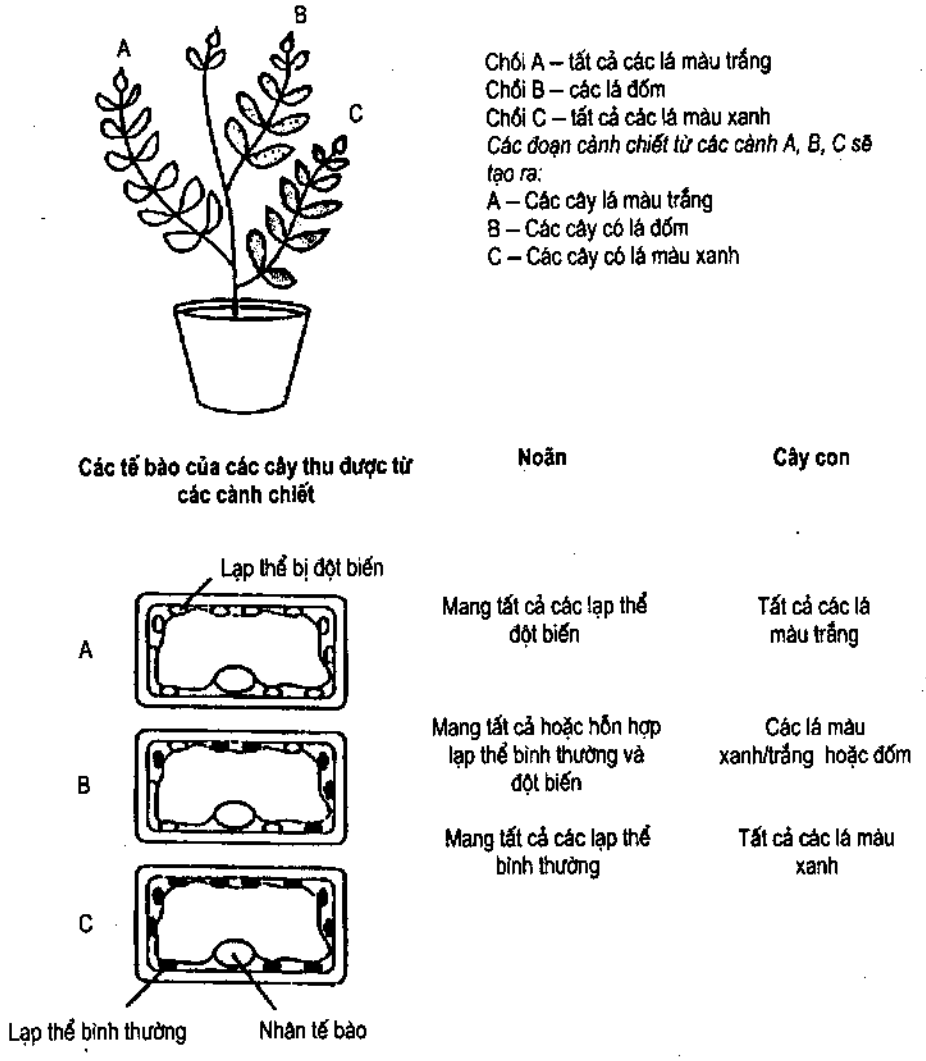
– Những cây có lá đốm trắng hoặc sọc trắng thường là do một số lạp thể bị đột biến mất khả năng tổng hợp chlorophyll. Sau mỗi lần phân bào, các lạp thể phân ly ngẫu nhiên về các tế bào con, và có thể một tế bào con nào đó chỉ được truyền các lạp thể bị đột biến. Các tế bào con của chúng sẽ không có chlorophyll và sinh ra các vệt trắng trên lá. Tần số của hiện tượng trên phụ thuộc vào số lạp thể trong tế bào và tỷ lệ giữa số lạp thể đột biến so với số lạp thể bình thường. Ở cây hoa Bốn giờ (*Mirabilia jalapa*) có thể cắt các cành xanh, trắng hoặc đốm rồi trồng lấy hoa và hạt. Các cành xanh, trắng, đốm có thể dùng để nhân giống vô tính (chẳng hạn bằng phương pháp chiết cành) tạo ra các dòng vô tính có kiểu hình lá có màu xanh, trắng đốm (*Hình 3.18*). Song, các hạt từ cây đốm có thể tạo ra hỗn hợp các cây xanh, trắng hoặc đốm. Điều đó là vì tế bào trứng của cây này có thể mang lạp thể bình thường, lạp thể đột biến hoặc hỗn hợp của cả hai loại.

– Một ví dụ khác về hiện tượng di truyền tế bào chất có ý nghĩa thương mại quan trọng ở thực vật là hiện tượng bất dục đực bào chất ở ngô. Các gen đột biến từ tế bào chất của mẹ ngăn cản sự hình thành hạt phấn. Hiện tượng đó có thể được sử dụng để sản xuất hạt lai F1. Khi

trồng xen kẽ dòng cây bất dục đực với dòng bình thường thì toàn bộ hạt trên dòng cây bất dục đực phải là hạt lai.

– Hiện tượng di truyền tế bào chất còn tìm thấy ở nấm. Một số đột biến ADN ty thể của nấm có liên quan đến hiện tượng giảm tốc độ sinh trưởng. Ở nấm men *Saccharomyces serevisiae*, những đột biến này được gọi là các đột biến "petite". Các dạng đột biến petite thường biểu hiện như kiểu hình lặn, vì khi nấm men kiểu dại lai với nấm men petite loại này sẽ sinh ra các bào tử có tốc độ sinh trưởng bình thường. Điều đó là do mỗi tế bào chứa nhiều ty thể, nên con lai, dù chứa hỗn hợp các ty thể bình thường và đột biến, khó có khả năng chỉ truyền các ty thể đột biến cho một tế bào con nào đó. Một số thể đột biến petite khác, chẳng hạn các đột biến petite ức chế, không phải là đột biến lặn. Khi lai những giữa những dòng nấm men mang những đột biến này với các dòng kiểu dại, đời con sinh ra có cả kiểu hình bình thường và kiểu hình đột biến petite, nhưng không tuân theo tỷ lệ phân ly của các quy luật Mendel.

Ở một số tế bào động vật nuôi cấy, đột biến ADN ty thể có thể dẫn đến tính kháng kháng sinh. Ở người, đột biến hệ gen ty thể là nguyên nhân gây nên một số bệnh di truyền theo dòng mẹ.



Hình 3.18. **Sự di truyền đặc điểm lạp thể ở các cây có lá đốm**

Sự di truyền theo dòng mẹ đối với quá trình tiến hóa

Trình tự ADN ty thể rất hữu ích để nghiên cứu tiến hóa, đặc biệt là nghiên cứu tiến hoá ở người, các loài linh trưởng và các loài động vật. Có một số lý do để giải thích cho điều này. Chẳng hạn ở động vật, tốc độ đột biến ADN ty thể thường cao hơn ADN nhân, nhưng trong phạm vi các tế bào của một cơ thể, hệ gen ty thể thường không có biến dị. Điều đó là vì chỉ có một số lượng nhỏ ty thể tham gia hình thành noãn bào. Khi đó, kiểu di truyền theo dòng mẹ đảm bảo cho sự đồng nhất các gen ty thể trong các tế bào soma của mỗi cá thể con. Ở thực vật, hệ gen ty thể có tập tính khác và có xu hướng tiến hoá chậm.

So với hệ gen nhân, khả năng đột biến nhanh hơn của hệ gen ty thể cùng với việc không có tái tổ hợp (do không có trao đổi chéo giữa các NST tương đồng như hệ gen nhân) cho phép các nhà nghiên cứu có thể dùng các biến đổi trên hệ gen ADN ty thể để ước lượng thời gian phân ly của các quần thể và các loài trong quá trình tiến hóa. Sở dĩ như vậy vì người ta có thể giả thiết rằng, hầu hết các biến đổi trong hệ gen ty thể là do các đột biến sinh ra, còn tần số xuất hiện đột biến thì tỷ lệ tuyến tính với thời gian tiến hóa (đây là cơ sở thiết lập "đồng hồ phân tử" trong các nghiên cứu tiến hóa). Dựa theo nguyên tắc đó, một số nghiên cứu đã tiến hành phân tích tính đa dạng trình tự ADN ty thể trong các quần thể người hiện nay và đưa đến kết luận rằng, tất cả các chủng người hiện nay đều là con cháu của một người phụ nữ duy nhất sống cách đây khoảng 200.000 - 300.000 năm. Tuy vậy, kết luận này vẫn đang còn tranh cãi do có bằng chứng cho thấy, tần số và tốc độ đột biến không phải luôn luôn ổn định trong suốt quá trình tiến hóa và còn phụ thuộc vào một số yếu tố, chẳng hạn như kích thước quần thể.

Hiệu ứng dòng mẹ

Có thể có sự nhầm lẫn giữa hiện tượng di truyền theo dòng mẹ với một số tình huống mà trong đó một số kiểu hình của con bị tác động bởi mẹ làm cho kiểu gen của con bị che khuất. Hiện tượng này gọi là hiệu ứng dòng mẹ. Hiệu ứng dòng mẹ xuất hiện ở nhiều loài sinh vật với mức độ khác nhau tùy theo sự đóng góp các thành phần protein và/hoặc ARN của mẹ vào phôi và sự phát triển của hợp tử. Một ví dụ là tính trạng chiều xoắn vỏ ốc ở loài ốc nước ngọt *Limnae*. Ở loài ốc này, một gen gồm hai alen quy định chiều xoắn vỏ ốc, trong đó alen *D* quy định chiều xoắn phải, còn alen *d* quy định chiều xoắn trái. Ốc nước ngọt là một loài lưỡng tính nên có thể tự phối hoặc giao phối. Khi ốc xoắn phải có kiểu gen *DD* lai với ốc xoắn trái *dd* thì toàn bộ ốc con có kiểu hình xoắn phải giống mẹ. Khi cho các cá thể F1 tự phối thì toàn bộ đời con F2 xoắn phải. Tiếp tục cho F2 tự phối thì thu được F3 có tỷ lệ 3 xoắn phải : 1 xoắn trái. Như vậy, tỷ lệ Mendel xuất hiện muộn một thế hệ so với thông thường. Cho dù hiện tượng này có hơi giống hiện tượng di truyền theo dòng mẹ, nhưng có một điều khác rõ ràng là kiểu hình ở đời con không luôn luôn giống mẹ.

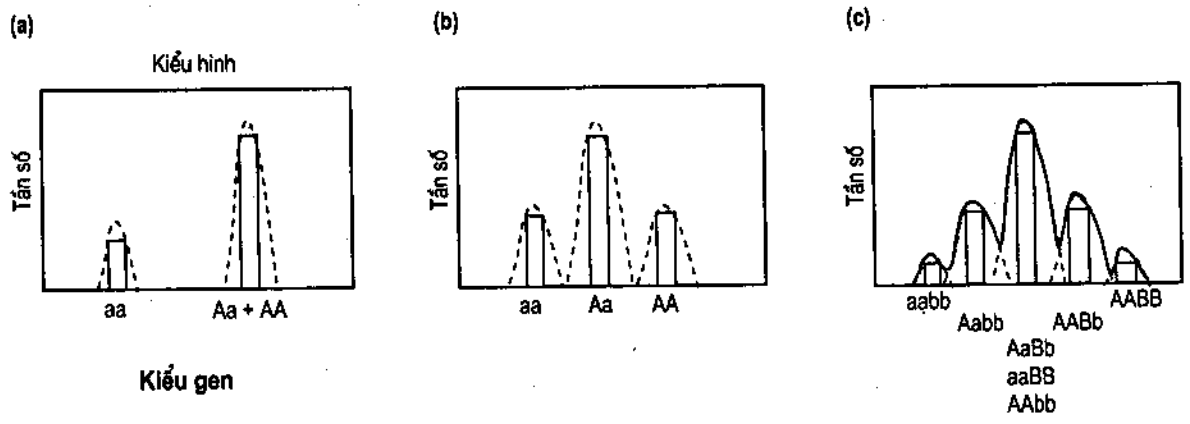
Vì sao tỷ lệ phân ly đúng lại xuất hiện muộn một thế hệ? Sự phát triển của ốc con do hệ gen mẹ quy định sẵn và có tác động đến chiều xoắn của vỏ ốc. Vì vậy, kiểu xoắn của ốc con phản ánh kiểu gen của mẹ chứ không phải kiểu gen của riêng nó. Ngoài ví dụ này, còn có một số ví dụ khác ở côn trùng: kiểu hình của con non thế hệ sau giống với mẹ, nhưng khi tuổi cá thể tăng lên hiện tượng đó dần dần mất đi.

III.G SỰ DI TRUYỀN CÁC TÍNH TRẠNG SỐ LƯỢNG

Tính trạng số lượng

Các tính trạng Mendel "truyền thống" thường do một gen có hai alen kiểm soát biểu hiện bằng hai hoặc ba kiểu hình khác biệt. Nếu một alen hoạt động chức năng là đủ thì kiểu hình trội sẽ được biểu hiện ở thể đồng hợp tử và dị hợp tử, còn kiểu hình lặn sẽ chỉ xuất hiện ở thể đồng hợp tử về alen không hoạt động chức năng (hoặc hoạt động chức năng yếu, *Hình 3.19a*). Nếu hoạt động chức năng của một alen không đủ mạnh thì thể dị hợp tử sẽ biểu hiện kiểu hình trung gian, nhưng vẫn có sự khác biệt với thể đồng hợp tử. Chẳng hạn, khi cho cây có hoa đỏ và trắng đồng hợp tử lai với nhau thu được các cá thể đời con có hoa màu hồng (*Hình 3.19b*). Trong cả hai trường hợp, mỗi alen đều có đóng góp vào sự biểu hiện kiểu hình. Trong những trường hợp này, môi trường chỉ gây ra những khác biệt nhỏ giữa các cá thể có cùng kiểu gen. Những biến dị này được gọi là biến dị **không liên tục**. Phần lớn các gen được quan tâm nghiên cứu thuộc nhóm này. Chúng có vai trò mang tính quyết định đối với sự biểu hiện của tính trạng.

Tuy vậy, không phải tất cả các gen đều có hiệu ứng rõ rệt như vậy. Thay vào đó, chúng có thể chỉ đóng góp một phần vào sự hình thành tính trạng. Ví dụ như chiều cao ở người thường không phải do một gen duy nhất kiểm soát (mặc dù ở người có tồn tại alen trội quy định tính trạng "lùn"). Phần lớn người trưởng thành có chiều cao trong khoảng 125 - 200cm. Nhưng trong khoảng chiều cao này, ở mọi kích thước đều có thể tìm được người có chiều cao tương ứng.



Hình 3.19. **(a) Phép lai một gen với hai alen có quan hệ trội hoàn toàn cho hai nhóm kiểu hình khác biệt phân ly theo tỷ lệ 3:1 (độ rộng của biểu đồ biểu diễn mức độ tác động của môi trường). (b) Nếu hai alen tác động cộng gộp thì dị hợp tử có kiểu hình trung gian và các nhóm kiểu hình ít khác biệt hơn. (c) Hai gen tác động cộng gộp sinh ra 5 nhóm kiểu hình với tỷ lệ 1:4:6:4:1. Tác động bổ sung của môi trường sẽ sinh ra biến dị liên tục.**

Khi có hai gen kiểm soát một tính trạng với vai trò ngang nhau, giả sử mỗi gen có hai alen và không có hiện tượng trội hoàn toàn giữa các alen, thì ở F2 sẽ có 5 nhóm kiểu hình với tỷ lệ phân ly 1:4:6:4:1 (*Hình 3.19c*). Với tác động bổ sung của các yếu tố môi trường, các nhóm kiểu hình này sẽ phủ lên nhau, tạo ra dãy **biến dị liên tục**. Khi tính trạng được quy định bởi sự tham gia của nhiều gen hơn, các đỉnh của đồ thị phân bố các nhóm kiểu hình sẽ được nối liền thành đường cong có một đỉnh duy nhất. Các tính trạng như vậy được gọi là các **tính**

trạng số lượng, vì lúc đó phải dùng phương pháp đo để xác định đặc điểm tính trạng của từng cá thể. Trong ví dụ này, chiều cao ở người là một tính trạng được kiểm soát di truyền rõ ràng. Con cái thường có xu hướng cao giống bố, mẹ. Một số chủng người cao, trong khi một số chủng người khác rất thấp. Tuy vậy, tất nhiên dinh dưỡng và một số yếu tố môi trường và tập tính cũng có ảnh hưởng đến chiều cao, ví như trẻ thiếu dinh dưỡng không thể cao được như anh, chị, em của nó vốn được nuôi dưỡng đầy đủ hơn.

Môi trường cũng có vai trò quan trọng trong sự hình thành nhiều tính trạng. Sự thiếu dinh dưỡng kinh niên hoặc đói ăn (thiếu protein hoặc vitamin) ở giai đoạn trẻ đang phát triển làm giảm sự tăng trưởng chiều cao dưới mức tiềm năng di truyền. Những tính trạng biến dị liên tục khác ở người có thể kể đến, gồm màu da, xu hướng bị bệnh tim hoặc tiểu đường, trí thông minh và tính cách. Các tính trạng đó còn được gọi là các tính trạng **đa nhân tố** hoặc **đa gen** vì có nhiều gen tham gia kiểm soát tính trạng. Các tính trạng biến dị liên tục bao gồm sản lượng và chất lượng các sản phẩm nông nghiệp (như sữa, thịt, hoa quả và hạt). Điều quan trọng cần biết là mức độ các gen và môi trường đóng góp như thế nào vào sự biểu hiện kiểu hình, nhằm có thể tối ưu hoá sản phẩm thông qua việc chọn lọc, chọn giống, nhân giống nhằm cải tiến hệ gen và thay đổi kỹ thuật canh tác, cải tạo môi trường.

Cơ sở hóa sinh học của hiện tượng di truyền đa gen có thể thấy được ở tính trạng màu hạt lúa mỳ. Một số dòng lúa mỳ có hạt màu trắng, trong khi một số dòng khác có màu đỏ. Mỗi alen hoạt động chức năng được phiên mã thành mARN rồi được dịch mã thành protein, đóng góp vào sự hình thành sắc tố đỏ. Càng nhiều alen hoạt động chức năng thì càng nhiều enzym và sắc tố được tạo ra, vì thế mà các alen đó được coi là có tác động cộng gộp. Màu sắc hạt lúa mỳ do ba locut gen kiểm soát, mỗi locut có hai alen. Phép lai bố, mẹ có kiểu gen *AABBCC* × *aabbcc* sẽ tạo ra 7 nhóm kiểu hình, lần lượt mang 0, 1, 2, 3, 4, 5 hoặc cả 6 alen trội ở F2, với tỷ lệ phân ly kiểu hình là 1 : 6 : 15 : 20 : 15 : 6 : 1. Chỉ có 1/64 số cá thể có kiểu gen *AABBCC* có hạt màu đỏ sẫm và 1/64 số cá thể có kiểu gen *aabbcc* có hạt màu trắng là giống bố, mẹ. Khung punnett cho phép lai này được biểu hiện trên *Hình 3.20*. Hiệu ứng của các nhân tố môi trường như nhiệt độ (làm biến đổi tốc độ phản ứng do enzym xúc tác), ánh sáng mặt trời và dinh dưỡng (làm biến đổi hàm lượng enzym và cơ chất tham gia phản ứng) gây nên sự biến động về màu sắc từng hạt đối với mỗi kiểu gen, làm giảm sự sai khác về kiểu hình giữa các nhóm kiểu gen liền kề.

	ABC	*ABc*	*AbC*	*aBC*	*Abc*	*aBc*	*abC*	*abc*
ABC	6	5	5	5	4	4	4	3
Abc	5	4	4	4	3	3	3	2
AbC	5	4	4	4	3	3	3	2
aBC	5	4	4	4	3	3	3	2
Abc	4	3	3	3	2	2	2	1
aBc	4	3	3	3	2	2	2	1
abC	4	3	3	3	2	2	2	1
abc	3	2	2	2	1	1	1	0

Hình 3.20. **Khung punnett cho thấy 7 nhóm kiểu hình về màu sắc hạt lúa mỳ do ba locut gen tác động cộng gộp tạo nên**

Cách đo lường biến dị liên tục

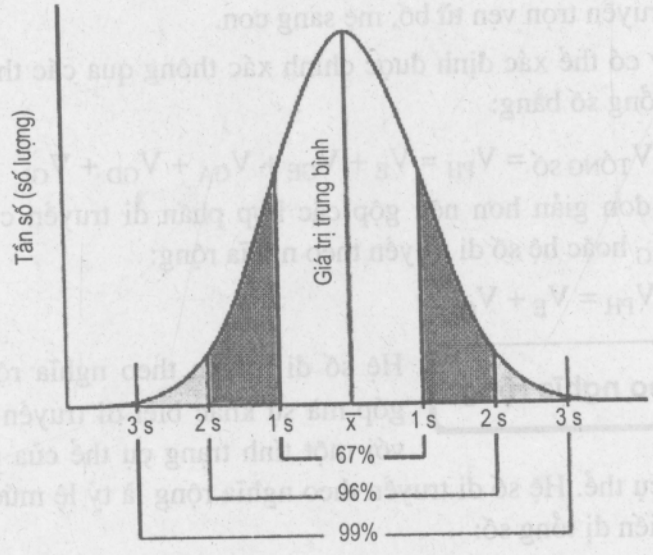
Hàm phân phối của một tính trạng biến dị liên tục, chẳng hạn như chiều cao ở người thường được biểu diễn bằng giá trị trung bình và phương sai của chúng. Giá trị trung bình được tính bằng tổng (Σ) các giá trị chiều cao (x_i) của tất cả các cá thể đo được chia cho tổng số cá thể (n). Giá trị trung bình thường được biểu diễn là x.

$$\text{Giá trị trung bình} = \bar{x} = \frac{\sum(x_i)}{n}$$

Sự biến động của phân phối quanh giá trị trung bình được gọi là phương sai (ký hiệu là s^2 hoặc V), được tính bằng tổng bình phương độ lệch giữa mỗi giá trị đo được (x_i) và giá trị trung bình $\bar{x}$, chia cho (n - 1).

$$s^2 = V = \frac{\sum(x_i - \bar{x})^2}{n-1}$$

Phương sai là một giá trị bình phương nên luôn có giá trị dương, đồng thời đơn vị của nó cũng được biểu diễn dưới dạng bình phương (như cm^2, kg^2, ...). Giá trị căn bậc hai của phương sai (s), được gọi là sai số tiêu chuẩn, cũng thường được dùng thay thế phương sai để khôi phục lại đơn vị biểu diễn của giá trị đo. Nhiều hệ thống sinh học thường biến động theo hàm phân bố chuẩn, là hàm toán học khi biểu diễn có đường cong dạng "hình chuông" (*Hình 3.21*). Đây là hàm phân phối cơ bản trong phần lớn các phép thống kê số học, nhưng không phải đúng trong mọi trường hợp. Nếu số đo một tính trạng nào đó tuân theo hàm phân phối chuẩn, thì có thể tính toán được vùng giá trị mong đợi. Chẳng hạn, vùng [giá trị trung bình ± s] chứa 68,3% tổng số cá thể, hoặc như vùng [giá trị trung bình ± 1,96s] chiếm 95% tổng số các cá thể được đo.



Hình 3.21. **Đồ thị biểu diễn hàm phân phối chuẩn, được xác định bởi giá trị trung bình và độ rộng của đường cong (tương ứng với sai số tiêu chuẩn)**

Cũng có những tính trạng có kiểu hình biến dị không liên tục nhưng lại được kiểm soát theo kiểu các tính trạng số lượng do có sự tham gia của nhiều gen và sự tác động của môi trường. Chẳng hạn như bệnh tiểu đường và ung thư. Ở những bệnh này, có sự tồn tại của

những alen gây bệnh, nhưng với sự kết hợp giữa kiểu gen với các yếu tố môi trường cùng với những sự kiện ngẫu nhiên khác mới đẩy một số người vượt ngưỡng từ trạng thái kiểu hình này sang kiểu hình khác (cụ thể là từ trạng thái bình thường sang trạng thái bệnh lý).

Các hợp phần phương sai

Phương sai kiểu hình tổng số V_{PH} bằng tổng các phương sai do các nhân tố riêng biệt tạo nên. **Phương sai do môi trường** (V_E) không phải là thành phần di truyền mà do môi trường sống không tuyệt đối giống nhau của mỗi cá thể gây nên. Một hợp phần khác là **phương sai tương tác giữa kiểu gen và môi trường** (V_{GE}). Ví dụ, một số người nhanh béo hơn những người khác khi được ăn uống tốt, hoặc nhanh gầy hơn khi hiếm thức ăn. Có những yếu tố môi trường khó đo lường, nhưng lại có ý nghĩa quan trọng. Chẳng hạn, một số giống lúa mỳ tăng sản lượng rõ rệt khi được cung cấp nhiều phân đạm, trong khi ở một số giống khác sản lượng không tăng lên rõ rệt khi bón nhiều phân đạm, nhưng những giống này lại có thể sống tốt hơn ở những môi trường nghèo dinh dưỡng.

Có ba hợp phần đúng là **các hợp phần di truyền**. Đó là **phương sai di truyền cộng gộp** (V_A) do tương tác cộng gộp của các alen (không phải là quan hệ trội lặn), vì thế mà kiểu gen *Aa* có kiểu hình nằm giữa các kiểu gen *AA* và *aa*. **Hợp phần di truyền trội** (V_D) do các alen trội tạo nên. Dị hợp tử biểu hiện phương sai này nhưng cũng truyền alen lặn cho đời con với xác suất giống như truyền alen trội. Sự biểu hiện của hợp phần này phụ thuộc vào alen nhận được từ bố, mẹ nên biến dị này không thật sự tương quan giữa bố, mẹ và con cái. Hợp phần di truyền thứ ba là **hợp phần tương tác gen** (V_{GI}) giữa các locut khác nhau (chẳng hạn tương tác át chế). Ví dụ: kiểu gen *AA* có thể biểu hiện khác nhau khi có mặt các alen *BB*, *Bb* hoặc *bb* ở một locut khác. Dường như các gen này được di truyền độc lập nên sự tổ hợp ở bố, mẹ sẽ bị phá hủy ở giai đoạn phát sinh giao tử. Những tổ hợp mới sẽ được tạo ra ở đời con nên sự tương tác sẽ không được di truyền trọn vẹn từ bố, mẹ sang con.

Các hợp phần này có thể xác định được chính xác thông qua các thí nghiệm nhân giống có kiểm soát. Biến dị tổng số bằng:

$$V_{TỔNG\ SỐ} = V_{PH} = V_E + V_{GE} + V_{GA} + V_{GD} + V_{GI}$$

Có thể biểu diễn đơn giản hơn nếu gộp các hợp phần di truyền của phương sai thành phương sai di truyền V_G hoặc hệ số di truyền theo nghĩa rộng:

$$V_{PH} = V_E + V_G$$

Hệ số di truyền theo nghĩa rộng

Hệ số di truyền theo nghĩa rộng H^2 là mức đóng góp mà sự khác biệt di truyền tạo ra ở biến dị đối với một tính trạng cụ thể của một quần thể cụ thể trong một môi trường cụ thể. Hệ số di truyền theo nghĩa rộng là tỷ lệ mức đóng góp di truyền đối với biến dị so với biến dị tổng số:

$$H^2 = \frac{V_G}{V_{PH}} = \frac{V_G}{V_E + V_G}$$

Hệ số di truyền theo nghĩa hẹp

Hệ số di truyền theo nghĩa hẹp h^2 là thước đo hữu ích đối với các nhà chọn giống động, thực vật. Nó là hợp phần phương sai di truyền cộng gộp V_{GA} (hoặc được

viết giản lược là V_A) chắc chắn được truyền cho đời con, cho phép tiên đoán khả năng đáp ứng đối với chương trình chọn lọc nhân tạo. Ở ví dụ màu sắc hạt lúa mỳ ở trên, số alen quy định màu đỏ cho phép tiên đoán màu sắc hạt trong thí nghiệm:

$$h^2 = \frac{V_A}{V_{PH}}$$

Kết quả sẽ rất khác biệt khi cả hiện tượng trội và tương tác cùng tác động lên kiểu hình vì hai cơ chế này không cùng được truyền trọn vẹn cho đời con.

Cách đo lường hệ số di truyền

(i) Chọn giống có kiểm soát

Các dòng nội phối sẽ trở thành đồng hợp tử nên sẽ không có biến dị do di truyền. Khi đó biến dị kiểu hình V_{PH} bằng V_E. Nếu lai hai dòng nội phối như vậy (P1 và P2) với nhau thì F1 đồng nhất về mặt di truyền và $V_{PH} = V_E$. Có ba cách ước tính V_E:

$$V_E = \frac{V_{E.P1} + V_{E.P2} + V_{E.F1}}{3}$$

F2 sẽ có đủ các kiểu gen và sẽ thể hiện tất cả các hợp phần di truyền của biến dị (V_G). Có thể tính các giá trị này từ $V_{PH.F2}$:

$$V_G = V_{PH.F2} - V_E$$

$$h^2 = \frac{V_{PH.F2} - V_E}{V_{PH.F2}}$$

Nếu toàn bộ hiệu ứng di truyền là do tương tác cộng gộp thì giá trị trung bình của F1 và F2 phải nằm giữa giá trị trung bình của các dòng bố, mẹ. Bất kỳ sự sai lệch nào cũng cho thấy đã có hiệu ứng trội hoặc tương tác át chế xảy ra. Phương sai di truyền của F2 (V_G) có thể phân thành V_{GA}, V_{GD} và V_{GI}, cho dù các hợp phần này có thể không có mặt đầy đủ trong từng trường hợp cụ thể. Có thể lai F1 trở lại với các dòng bố, mẹ để tạo ra hai dòng lai ngược (B1 và B2). Có thể tính được mức đóng góp của hiệu ứng trội và hiệu ứng cộng gộp đối với phương sai trong mỗi dòng (F2, B1 và B2) để từ đó tính V_A và V_D.

(ii) Hệ số di truyền thực nghiệm

Có thể tính hệ số di truyền từ kết quả chọn lọc. Giả sử một quần thể cà chua có khối lượng quả trung bình (M_T) là 30g. Người ta chọn những quả to, có khối lượng trung bình (M_P) là 40g để nhân giống. Phân sai chọn lọc trong trường hợp này ($M_P - M_T$) là 10g. Nếu thế hệ sau có khối lượng quả trung bình (M_O) là 34g thì:

$$\text{Hệ số di truyền} = \frac{\text{Giá trị gia tăng}}{\text{Phân sai chọn lọc}}$$

$$H = \frac{M_o - M_T}{M_P - M_T} = \frac{34-30}{40-30} = \frac{4}{10} = 0,4$$

(iii) Hồi quy con cái - bố, mẹ

Có thể tính hệ số di truyền bằng cách so sánh các cá thể có quan hệ họ hàng (đây là phương pháp duy nhất áp dụng cho người) và có thể ước tính hệ số di truyền theo nghĩa hẹp từ

các so sánh đó bằng cách tính hệ số hồi quy "bố, mẹ - con cái". Nếu mọi biến dị đều do tác động cộng gộp thì kiểu hình của con sẽ nằm ở khoảng giữa kiểu hình của hai bố, mẹ. Vậy, nếu kiểu hình của con nằm ở khoảng giữa hai bố, mẹ thì đồ thị sẽ là một đường thẳng với độ dốc là 1 khi phương sai hoàn toàn do tác động cộng gộp của các alen. Đồ thị kiểu hình của một cặp bố, mẹ so với con, hoặc của anh, chị em với nhau sẽ có hệ số hồi quy 0,5 nếu hệ số di truyền bằng 1. Hiệu ứng môi trường sẽ làm sai lệch các điểm và làm giảm hệ số hồi quy. Khi phương sai hoàn toàn do môi trường thì không có sự tương ứng giữa bố, mẹ và con cái ($h^2 = 0$). Các điểm sẽ phân tán ngẫu nhiên trên đồ thị và độ nghiêng của đường hồi quy bằng 0. Công thức cho đồ thị đường thẳng hồi quy là $y = a + bx$, với b là độ nghiêng của đường hồi quy và đường hồi quy cho trường hợp con có kiểu hình nằm giữa bố, mẹ là:

$$h^2 = b$$

Giả thiết không có nghĩa là không có sự tương ứng và $b = 0$. Giá trị b được tính bằng phép phân tích hồi quy và mức độ sai khác thống kê được xác định bằng chuẩn Student. Nếu $b > 1$ thì chắc chắn con lai có ưu thế lai (hiện tượng ưu thế lai hoặc siêu trội), điều có thể xảy ra trong các chương trình chọn giống.

Phương pháp khác là tính hệ số tương quan r giữa giá trị trung bình của bố, mẹ và con cái. Nếu các giá trị trung bình của bố, mẹ và con cái có kiểu phân bố giống nhau (có cùng giá trị trung bình và phương sai) thì tính toán trên sẽ cho thấy giá trị h^2. Sự khác biệt chính giữa các cách tính này là người ta giả thiết rằng không có nguyên nhân nào và quan hệ hiệu ứng nào tác động đến hệ số tương quan, trong khi để phân tích hồi quy thì người ta giả thiết rằng kiểu gen của bố, mẹ có ảnh hưởng đến đời con.

Những hạn chế của hệ số di truyền

Hệ số di truyền có liên quan với biến dị tổng số. Nếu cây trồng (ví dụ Cà chua) được trồng trong điều kiện môi trường rất biến động thì môi trường sẽ sinh ra nhiều biến dị và hệ số di truyền sẽ thấp. Nếu cũng những cây cà chua đó được trồng trong các điều kiện môi trường ổn định hơn thì sẽ ít có biến dị và hệ số di truyền sẽ cao, dù cho chúng có cùng kiểu gen. Hệ số di truyền chỉ có ý nghĩa về mặt số lượng trong điều kiện cụ thể khi nó được đo. Nếu hệ số di truyền cao, quá trình chọn giống theo các nguyên tắc di truyền sẽ cải tiến được năng suất. Nếu hệ số di truyền thấp thì việc thay đổi và kiểm soát môi trường sẽ có hiệu quả hơn. Vì vậy, đối với những dòng đồng nhất về mặt di truyền, và ở những loài thực vật tự thụ phấn như lúa mỳ, trong điều kiện môi trường ổn định, biến dị hầu như không xuất hiện.

Những locut tính trạng số lượng

Có nhiều gen tham gia quy định tính trạng số lượng phân bố rải rác suốt hệ gen. Nơi định vị của chúng được gọi là các locut tính trạng số lượng (viết tắt là QTL - *q*uantitative *t*rait *l*oci). Việc xác định vị trí của chúng rất có ý nghĩa trong y học và trồng trọt. Đối với các tính trạng mẫn cảm với bệnh tật, việc xác định được các gen có liên quan là rất hữu ích, vì để từ đó xác định được chức năng bình thường của các gen và xây dựng phương pháp điều trị phù hợp. Đối với việc chọn giống và nhân giống ở động vật và thực vật, việc xác định được các cá thể non có các đặc điểm tốt mà không cần đợi đến khi chúng trưởng thành rất có ý nghĩa. Qua việc này, có thể loại bỏ sớm được những kiểu gen xấu, chọn lọc và phân tách được những kiểu gen có tiềm năng năng suất và chất lượng cao.

Các bước để chọn giống bao gồm chọn ra các dòng thuần khác biệt nhau về đặc điểm cần quan tâm và liên kết với các dấu chuẩn (thường là dấu chuẩn VNTR) ở vị trí các mẫu dò. Tiến hành lai chúng với nhau, rồi kiểm tra con lai ở thế hệ F2 và các thế hệ sau về tính trạng quan tâm cũng như các biến dị ở vị trí các mẫu dò. Nếu tính trạng có tương quan với sự di truyền của alen đánh dấu cụ thể nào đó thì có thể xác định được một gen hoặc nhiều gen định vị gần gen đánh dấu trên phân tử ADN có tác động đến tính trạng.

Cách tương tự cũng có thể thực hiện ở người, đặc biệt là các locut quy định sự mẫn cảm với bệnh tật. Tuy vậy, việc này phức tạp và khó khăn hơn vì số lượng người trong một gia đình thường ít. Trường hợp thông thường là có một hay hai gen gây ra hầu hết các biến dị, bên cạnh nhiều gen có hiệu ứng nhỏ. Khó phát hiện được các gen đóng góp khoảng 5% hoặc ít hơn cho biến dị của tính trạng. Một nhân tố di truyền quan trọng nhất ở bệnh tiểu đường phụ thuộc insulin là alen của vùng khởi động của gen mã hoá bản thân insulin, sau đó là một vài locut khác ít quan trọng hơn.

Những nghiên cứu trên người

Trẻ sinh đôi cùng trứng có cùng kiểu gen nên biến dị giữa chúng hầu hết do tác động của môi trường (điều đó không hoàn toàn đúng đối với các bé gái do sự bất hoạt ngẫu nhiên một trong hai nhiễm sắc thể X). Việc so sánh biến dị giữa các trẻ sinh đôi cùng trứng được nuôi trong cùng một môi trường cho ta biết hiệu ứng môi trường "trong bản thân trẻ sinh đôi", còn việc so sánh các trẻ sinh đôi được nuôi riêng rẽ trong các gia đình khác nhau cho ta biết hiệu ứng môi trường "giữa các gia đình" đối với biến dị. **Trẻ sinh đôi khác trứng** chỉ đơn giản là anh, chị, em ruột được sinh ra cùng một thời điểm. Sự khác biệt giữa các anh, chị, em ruột bình thường với anh, chị, em sinh đôi khác trứng chính là biến dị do được sinh ra ở những thời điểm riêng biệt. Các trẻ sinh đôi cùng được một gia đình nhận làm con nuôi hoặc được nhận nuôi riêng có thể có tương quan với nhau hoặc tương quan với bố, mẹ nuôi, hoặc bố, mẹ đẻ, với anh, chị, em nuôi hoặc anh, chị, em ruột, giúp nghiên cứu mức đóng góp của môi trường và các nhân tố di truyền đối với kiểu hình. Con đẻ của các bố, mẹ nuôi sẽ là đối chứng. Tương quan của con nuôi với giá trị trung bình kiểu hình của bố, mẹ đẻ là thước đo hệ số di truyền theo nghĩa hẹp.

Trên đây là cách ước tính hợp phần môi trường của biến dị kiểu hình, phần biến dị còn lại là do các sai khác di truyền, nghĩa là hệ số di truyền theo nghĩa rộng. Cần phải nhớ rằng các gia đình cũng chia sẻ chung môi trường, làm cho việc tách biệt hiệu ứng môi trường và di truyền trở nên khó khăn. Các con nuôi được dùng làm đối chứng cho mục đích này vì chúng cùng chia sẻ các gen, nhưng không cùng chia sẻ môi trường. Hiệu ứng môi trường được tính bắt đầu từ khi làm tổ trong dạ con và tiếp tục là sự phát triển trong tử cung. Trẻ sinh đôi và anh, chị, em ruột của nó có cùng mẹ nên chúng cùng chia sẻ hiệu ứng môi trường trước khi sinh cũng như các gen. Vậy hiệu số của hệ số di truyền theo nghĩa rộng và hệ số di truyền theo nghĩa hẹp (H^2 - h^2) có thể gán cho hiệu ứng trội hoặc tương tác gen.

Rõ ràng là việc ước tính hệ số di truyền cho các tính trạng ở người kém chính xác hơn ở động vật và thực vật, vì số mẫu nhỏ hơn và không thể là đại diện cho quần thể. Các số liệu đó còn có thể là các số liệu sưu tập từ các nguồn không chính thức.

Các tính trạng giống nhau có thể xuất hiện ở cả hai cặp trẻ sinh đôi, và có thể sử dụng để ước tính hệ số di truyền. Sự sai khác về tần số các tính trạng giống nhau giữa các trẻ sinh đôi

cùng trứng và khác trứng được lấy làm thước đo hoặc chỉ thị cho hợp phần di truyền của tính trạng. Tần số này ở trẻ sinh đôi cùng trứng cao hơn ở trẻ sinh đôi khác trứng đối với các đặc điểm như vui, buồn thất thường (80% : 20%), bệnh tâm thần phân liệt (80% : 13%) và huyết áp (63% : 36%), nhưng chỉ là 95% : 87% đối với bệnh sởi. Điều đó cho thấy tính biến dị di truyền cao đối với bệnh tâm thần phân liệt và tính biến dị di truyền thấp về tính mẫn cảm với bệnh sởi.

Các tính trạng số lượng ở người

Nhiều tính trạng ở người là các tính trạng số lượng. Ví dụ đơn giản là màu da. Các số liệu cho thấy có ít nhất bốn gen quan trọng (có thể là sáu gen hoặc hơn), với 8 - 12 alen tác động cộng gộp, tạo ra 9 - 13 lớp kiểu hình. Các tính trạng lý thú hơn là trí thông minh và tính cách. Sẽ có ích nếu biết được một cách tương đối vai trò của các yếu tố môi trường (gia đình và xã hội) và yếu tố di truyền đóng góp như thế nào vào biến dị của trí thông minh. Nếu hệ số di truyền của trí thông minh cao thì sự đầu tư tăng lên cho giáo dục sẽ không làm tăng số người tài năng. Những nhóm người có khả năng hoặc kém khả năng hơn sẽ cần những hình thức đào tạo khác nhau. Nếu hệ số di truyền thấp (hiệu ứng môi trường cao) thì việc chi phí tiền bạc để đưa môi trường đến cho những người kém khả năng ngang bằng với những người có khả năng sẽ có hiệu quả cao. Một số nghiên cứu gần đây xác nhận hệ số di truyền của IQ là khoảng 0,6, còn thường là xung quanh giá trị 0,8. Điều đó có ý nghĩa gì? Nên nhớ rằng, hệ số di truyền là tỷ lệ biến dị do di truyền so với biến dị tổng số. Nếu các thử nghiệm IQ được thực hiện ở nhóm các học sinh lớp giữa của trình độ phổ thông thì tất cả chúng đều có cơ hội như nhau khi các gia đình đều có mối quan tâm cao đến việc giáo dục trẻ, khi đó yếu tố môi trường có thể thấp. Nếu thực hiện trên toàn bộ dân cư của một thành phố, kể cả người nghèo lẫn người giàu, thì biến dị do môi trường sẽ lớn hơn nhiều, nên hệ số di truyền có thể thấp, trong khi thực tế là các gen đều như nhau trong cả hai cách thử. Một nhóm nghiên cứu phát hiện rằng: một alen mã hóa cho thụ thể của yếu tố sinh trưởng giống insulin-2 có tần số cao hơn (khoảng 30%) ở những trẻ năng khiếu, so với tần số này (khoảng 16%) ở các trẻ thông thường. Đáng chú ý rằng hầu như mọi người có alen này đều bình thường.

Tính cách và hành vi chắc chắn cũng là những tính trạng được điều khiển bởi yếu tố di truyền, bởi vì chúng phụ thuộc khá lớn vào hoạt động của các chất dẫn truyền thần kinh và các thụ thể của chúng. Điều đó giải thích vì sao các chất điều trị tâm thần lại có tác động đến trí tuệ. Đến nay, đã xác định được hơn 100 loại chất dẫn truyền thần kinh khác nhau, mỗi chất dẫn truyền thần kinh lại bao gồm một số thụ thể đặc trưng của nó. Nên chắc chắn có sự biến dị di truyền lớn liên quan đến những hoạt động dẫn truyền thần kinh. Gen mã hóa thụ thể serotonin có hai alen ở vùng khởi động. Một có hoạt tính cao còn một có hoạt tính thấp. Những người có alen hoạt tính cao hay biểu hiện lo lắng và băn khoăn hơn so với những người có alen hoạt tính thấp. Một số nghiên cứu so sánh các trẻ sinh đôi và anh, chị, em ruột cho thấy hệ số di truyền theo nghĩa rộng là 30 - 45% (và hệ số di truyền theo nghĩa hẹp là 10 - 20%) cho "tính dễ thương", "tận tâm", "hướng ngoại", "dễ bị kích động" và "cởi mở" bằng cách tính điểm từ các nghiên cứu điều tra về tâm lý học.

III.H SỰ XÁC ĐỊNH GIỚI TÍNH

Sự xác định giới tính

Giới tính của một cá thể sinh vật có thể được xác định và biểu hiện ở các mức độ khác nhau. Ở đây, chúng ta chủ yếu chỉ đề cập đến sự xác định giới tính sơ cấp, nghĩa là liên quan đến sự kiện xác định một cá thể trong quá trình phát triển sẽ phát triển tinh hoàn hay buồng trứng. Các đặc điểm giới tính thứ cấp (là các đặc điểm sinh trưởng phát triển liên quan chặt chẽ với giới tính này hay giới tính kia), chẳng hạn như màu lông, sừng, bờm, v.v... cũng được kiểm soát bởi cơ chế di truyền.

Các sinh vật đơn bào có thể có các hệ thống xác định giới tính đơn giản, trong khi các sinh vật đa bào khác nhau khá nhiều về cơ chế hình thành giao tử đực và giao tử cái (vốn cần thiết cho quá trình sinh sản hữu tính). Nhiều trường hợp một cá thể có cả cơ quan sinh sản đực và cái. Các cá thể đó được gọi là **lưỡng tính**. Hiện tượng này khá phổ biến ở nhiều loài động vật không xương sống và ở thực vật. Các sinh vật lưỡng tính có thể có đồng thời hai giới tính hoặc có thể chuyển từ giới tính này sang giới tính khác. Ở các loài này, rõ ràng các gen xác định giới tính dường như ít quan trọng hơn so với trường hợp các sinh vật bậc cao. Tuy vậy, nhiều bằng chứng cho thấy: ở một số loài động vật không xương sống lưỡng tính, các gen có thể tham gia điều hoà thời điểm hình thành từng giới tính và thời gian kéo dài của các pha giới tính.

Hiện tượng lưỡng tính không chỉ giới hạn ở động vật không xương sống. Một số loài cá như Cá vược có thể chuyển đổi giới tính do ảnh hưởng của môi trường hoặc hormon. Một số loài chim nuôi đôi khi có thể tự nhiên chuyển đổi giới tính. Ví dụ rõ ràng về tác động của môi trường đến sự hình thành giới tính là trường hợp ở cá sấu: nhiệt độ ấp trứng xác định giới tính của cá thể.

Các hệ thống xác định giới tính đơn giản

Có lẽ cơ chế xác định giới tính đơn giản nhất là ở nấm men. Hai alen của một gen xác định kiểu giới tính. Ở nấm men *Saccharomyces serevisiae*, một gen trên nhiễm sắc thể (NST) số 3 là gen *MAT* có hai alen *a* và α. Gần như suốt vòng đời của mình, *S. serevisiae* là đơn bội và tế bào nấm men chỉ mang alen *a* hoặc α. Các alen này xác định kiểu giới tính. Chỉ có các tế bào nấm men có kiểu giới tính ngược nhau mới có thể kết hợp để tạo thành tế bào lưỡng bội. Như vậy, có thể coi gen *MAT* là kiểu xác định giới tính ra đời sớm. Hệ thống xác định giới tính tương tự cũng được tìm thấy ở các loài như *Chlamydomonas*.

Côn trùng cánh màng (như kiến, ong và các côn trùng bay) thường có phương thức xác định giới tính không bình thường. Ong đực phát triển từ trứng không được thụ tinh và là đơn bội. Ong cái (gồm ong thợ và ong chúa) phát triển từ trứng được thụ tinh và là lưỡng bội. Người ta cho rằng, trong trường hợp này, mức bội thể là nhân tố xác định giới tính. Tuy nhiên, cơ chế xác định giới tính còn phụ thuộc vào một gen đa alen. Nếu gen này là dị hợp tử thì ong đó là ong cái. Ong đực đơn bội không thể là dị hợp tử vì mỗi gen chỉ có một bản sao. Nội phối qua nhiều thế hệ sẽ tạo ra mức đồng hợp tử cao và những dòng ong đực lưỡng bội thuần chủng cao đã được phát hiện.

Các hệ thống nhiễm sắc thể xác định giới tính

Ở nhiều loài, các gen xác định giới tính liên kết với các NST đặc biệt được gọi là các NST giới tính. Có một số hệ thống NST xác định giới tính, bao gồm các hệ thống sau:

+ **Hệ thống XX-XO**. Hệ thống này được tìm thấy ở nhiều loài côn trùng. Con cái có cặp nhiễm sắc thể XX. Con đực chỉ có một nhiễm sắc thể X. Trường hợp này gặp ở Châu chấu và Rệp *Protenor* (nên đôi khi nó được gọi là hệ thống Protenor).

+ **XX-XY**. Hệ thống này gặp ở động vật có vú và một số loài côn trùng, kể cả *Drosophila*. Giới cái có hai nhiễm sắc thể X, giới đực có một nhiễm sắc thể X và một nhiễm sắc thể Y.

+ **ZZ-ZW**. Đây là hệ thống đảo ngược của hệ thống XX-XY: giới cái là ZW, còn giới đực là ZZ. Hệ thống này được tìm thấy ở chim, bướm và rắn.

Các thuật ngữ đồng giao tử và dị giao tử được sử dụng để mô tả các hệ thống này. Đồng giao tử để chỉ giới tính chỉ sinh ra các giao tử mang các NST giới tính giống nhau. Ví dụ, với hệ thống XX-XY, giới cái là đồng giao tử vì chỉ sinh ra một loại giao tử mang nhiễm sắc thể X. Với hệ thống ZZ-ZW, giới cái lại là dị giao tử vì sinh ra hai loại giao tử, một mang Z, một mang W.

Cho dù hai loài có hệ thống NST xác định giới tính giống nhau thì cũng không có nghĩa là các gen xác định giới tính hoạt động theo cách giống nhau. Các ví dụ ở mục dưới sẽ minh hoạ cho luận điểm này.

Sự xác định giới tính ở *Drosophila*

Giới tính ở *Drosophila* được xác định bởi tỷ lệ các nhiễm sắc thể X so với số bộ nhiễm sắc thể thường (A). Khi tỷ lệ này bằng 1 hoặc lớn hơn thì đó là ruồi cái. Khi nó bằng hoặc nhỏ hơn 0,5 thì đó là ruồi đực. Nếu tỷ lệ đó nằm giữa 0,5 - 1 thì đó là ruồi lưỡng tính. Một số ví dụ được đưa ra ở *Bảng 3.9*.

Bảng 3.9. **Tỷ lệ NST X so với bộ các NST thường và sự xác định giới tính ở *Drosophila***

Số nhiễm sắc thể X	Số bộ nhiễm sắc thể thường (A)	Tỷ lệ X : A	Giới tính
3	2	1,50	Cái
3	3	1,00	Cái
2	2	1,00	Cái
2	3	0,67	Trung tính
1	2	0,50	Đực
1	3	0,33	Đực

Các tỷ lệ "cực đoan" như 0,33 và 1,50 sinh ra những con ruồi được gọi là nửa đực hoặc nửa cái. Những con ruồi này phát triển kém và thường chết sớm.

Thực tế là sự xác định giới tính là kết quả của sự cân bằng các nhiễm sắc thể X và các nhiễm sắc thể thường cho thấy các gen quy định giới cái tập trung trên nhiễm sắc thể X và các gen quy định giới đực trên nhiễm sắc thể thường. Có một điểm quan trọng cần lưu ý đối với nhiễm sắc thể Y. Số liệu trên đây cho thấy nó không có vai trò gì trong sự xác định giới tính ở

Drosophila. Điều đó đúng, nhưng dù những con ruồi không có nhiễm sắc thể Y vẫn là ruồi đực thì chúng vẫn bất thụ vì không có gen trên nhiễm sắc thể Y cần thiết để phát triển tinh trùng có khả năng thụ tinh.

Cơ chế cân bằng di truyền điều khiển sự xác định giới tính tương tự cũng có ở các loài khác như giun tròn *Ceanorhabditis elegans*. Tuy nhiên, cơ chế có hơi phức tạp hơn vì có cả con đực và các cá thể lưỡng tính tồn tại trong loài.

Một đặc điểm quan trọng của sự xác định giới tính ở ruồi giấm là có những con ruồi không bình thường được gọi là lưỡng tính. Chúng là kết quả của sự không phân ly NST ở các tế bào soma. Nếu kết quả đó dẫn đến sự thay đổi số lượng nhiễm sắc thể X trong tế bào thì tỷ lệ X : A sẽ thay đổi và có thể ảnh hưởng đến giới tính của tế bào. Điều đó có thể xảy ra vì ở ruồi giấm, giới tính được xác định riêng ở mỗi tế bào. Nếu tế bào đó tiếp tục phân chia, nó sẽ sinh ra dòng tế bào, và phụ thuộc vào vị trí của các tế bào đó trong cơ thể, chúng có thể biệt hoá tạo ra cấu trúc có giới tính ngược với giới tính của các phần khác của cơ thể. Trường hợp cực đoan nhất là việc mất nhiễm sắc thể X ở lần phân chia đầu tiên của hợp tử có thể sinh ra những con ruồi một nửa cơ thể là đực, một nửa là cái. Hiện tượng này không có ở động vật có vú khi các đặc điểm giới tính thứ cấp do các hormon giới tính kiểm soát.

Sự xác định giới tính ở người

Sự xác định giới tính ở người là đặc trưng nhất so với ở các loài động vật có vú khác. Tuy rằng các NST giới tính cũng là XX và XY nhưng cơ sở di truyền của sự xác định giới tính khác đáng kể so với ở *Drosophila*. Những hiểu biết của chúng ta xuất phát từ những nghiên cứu các thể dị bội NST giới tính. Hiện tượng dị bội thể NST giới tính phát sinh thường xuyên hơn dị bội thể NST thường vì chỉ có ít gen trên nhiễm sắc thể Y và vì hiện tượng bất hoạt nhiễm sắc thể X nên chỉ có một nhiễm sắc thể X biểu hiện ở mọi tế bào. Cho nên sự thay đổi số lượng nhiễm sắc thể X ít ảnh hưởng đến khả năng sống sót hơn so với sự thay đổi trên các NST thường. Giới tính của một số thể dị bội NST giới tính được trình bày ở *Bảng 3.10*. Hầu như tất cả các thể dị bội nhiễm sắc thể giới tính đều bất thụ và các cá thể có sự sai lệch càng lớn so với mức bình thường càng có trí tuệ kém.

Bảng 3.10. **Tương quan giữa số lượng NST X và sự xác định giới tính ở người**

Nhiễm sắc thể giới tính	Số lượng nhiễm sắc thể	Giới tính
X	45	Nữ
XXX	47	Nữ
XXXX	48	Nữ
XXXXX	49	Nữ
XYY	47	Nam
XXY	47	Nam
XXXY	48	Nam

Từ những ví dụ này, chúng ta thấy rõ là ở mức độ nhiễm sắc thể, sự có mặt của nhiễm sắc thể Y là nhân tố xác định giới tính đực ở người. Ở giai đoạn phôi sớm, sự có mặt của nhiễm sắc thể Y làm cho tuyến sinh dục chưa biệt hoá phát triển nhanh chóng hơn để thành tinh hoàn. Ở chim, giới cái là ZW, sự có mặt của nhiễm sắc thể W cảm ứng sự phát triển buồng

trứng từ tuyến sinh dục chưa biệt hoá. Một gen đặc biệt có cả ở người và chuột, gen SRY (hay TDF) nằm trên nhiễm sắc thể Y đã được phân lập, chịu trách nhiệm chuyển cái thành đực ở giai đoạn phôi.

Sự xác định giới tính ở thực vật

Hầu hết thực vật hạt kín là lưỡng tính, hoa mang cả cơ quan sinh dục đực và cái. Tuy nhiên, ở một số loài, hoa đực và hoa cái được sinh ra ở các cây khác nhau (những loài này gọi là thực vật khác gốc). Một chi thực vật mà nguyên lý xác định giới tính được xác định bởi nhiễm sắc thể là chi *Melandrrium*. Ở đây, tỷ lệ X : Y là nhân tố quan trọng quyết định một cây hình thành hoa đực hay hoa cái. Nghĩa là, các gen trên cả hai nhiễm sắc thể giới tính tương tác để xác định giới tính. Các nghiên cứu về những thực vật bị đột biến mất đoạn trên nhiễm sắc thể X hoặc Y do chiếu xạ cho thấy nhiễm sắc thể Y có chứa những vùng ức chế sự phát triển giới tính cái và cảm ứng sự phát triển giới đực, trong khi nhiễm sắc thể X có những vùng kích thích sự phát triển hoa cái.

Các đặc điểm giới tính thứ cấp

Sự hình thành các đặc điểm giới tính thứ cấp cũng chịu sự kiểm soát di truyền. Có thể minh hoạ bằng gen *Tfm* ở động vật có vú. Gen này mã hoá protein hoạt động như thụ quan hormon steroid tính đực, hormon testosteron và biểu hiện cả ở giới đực và cái. Testosteron chỉ được sản sinh ra ở tinh hoàn và chịu trách nhiệm cho sự hình thành các đặc điểm giới tính thứ cấp ở giới đực. Các đột biến của gen này sinh ra hội chứng được gọi là hội chứng cái hoá tinh hoàn. Ở đây, sự có mặt của nhiễm sắc thể Y làm cho tinh hoàn được hình thành và sinh ra testosteron. Tuy nhiên, hormon không tác động đến các tế bào đích vì chúng không có thụ quan hoạt động chức năng. Các cá thể mắc hội chứng này phát triển như con cái bất thụ. Cơ quan sinh dục của chúng là tinh hoàn nhưng nằm sâu bên trong và NST giới tính của chúng là XY.

Một sai lệch trong biệt hoá giới tính sơ cấp và thứ cấp khác xảy ra ở các con gia súc cái bất toàn. Chúng có thể là cừu, dê hoặc gia súc lớn có sừng phát triển như những con cái bất thụ nhưng lại có tinh hoàn khiếm khuyết nằm sâu bên trong. Các tế bào tinh hoàn và các cơ quan khác đều có hai nhiễm sắc thể X nhưng máu lại có những tế bào có cả nhiễm sắc thể X và Y. Sự phát triển không bình thường này chỉ thấy có ở những con cái là một trong hai con sinh đôi khác giới. Máu cung cấp cho hai phôi hòa hợp lại ở dạ con nên các tế bào máu XY và hormon có thể đi vào vòng tuần hoàn của con cái. Điều đó buộc cơ quan sinh dục phát triển thành tinh hoàn và hạn chế một phần sự phát triển bình thường của con cái cho dù nó có kiểu gen XX. Con sinh đôi là con đực không bị tác động bởi sự có mặt của các tế bào máu XX.

Tiến hóa của nhiễm sắc thể giới tính

Có những bằng chứng cho thấy NST giới tính tiến hoá từ cặp NST thường. Quá trình này phải xảy ra từ rất lâu. Ví dụ như ở rắn, NST giới tính đã được nghiên cứu ở các loài rắn nguyên thuỷ và các loài tiến hoá cao. Ở các loài nguyên thuỷ, hai NST giới tính giống hệt nhau. Hai NST này chỉ có khác biệt là nhiễm sắc thể W sao chép ở pha S muộn hơn. Ở những loài tiến hoá cao hơn, nhiễm sắc thể W giảm kích thước, trở thành dị nhiễm sắc và khác biệt hẳn so với nhiễm sắc thể Z. Cặp nhiễm sắc thể giới tính chỉ duy trì một vùng tương đồng nhỏ. Vùng đó được gọi là vùng autosom giả (hay NST thường

giả) và chỉ cần cho việc hai nhiễm sắc thể này có thể tiếp hợp và phân ly chính xác trong quá trình giảm phân. Ở động vật có xương sống, NST giới tính chỉ có mặt ở giới dị giao tử (Y hoặc W), nói chung chúng thường chỉ mang một số ít gen và tích luỹ một lượng lớn trình tự ADN vệ tinh và chất dị nhiễm sắc cơ định.

III.I SỰ DI TRUYỀN LIÊN KẾT GIỚI TÍNH

Sự di truyền liên kết giới tính

Ở động vật có vú, vì con đực chỉ có một nhiễm sắc thể X còn con cái có hai nhiễm sắc thể X nên các gen trên nhiễm sắc thể này có kiểu di truyền khác biệt với di truyền Mendel. Những tính trạng do các gen đặc thù giới tính quy định được gọi là **sự di truyền liên kết giới tính.**

Các alen lặn không được biểu hiện ở con cái dị hợp tử. Con đực có một nhiễm sắc thể X (còn gọi là **bán hợp tử**) nên kiểu hình alen lặn được biểu hiện. Một ví dụ về tính trạng này là bệnh mù màu (đỏ và xanh lục) ở người, do các alen lặn trên nhiễm sắc thể X quy định. Alen bình thường được ký hiệu là *CB*, còn alen đột biến gây bệnh mù màu là *cb*. Các kiểu gen của bệnh này được trình bày ở *Bảng 3.11*.

Bảng 3.11. **Các kiểu gen có thể có của locut quy định bệnh mù màu**

Nam bình thường	*CB/Y*
Nam mù màu	*cb/Y*
Nữ mù màu	*cb/cb*
Nữ bình thường	*CB/CB* hoặc *CB/cb*

Người phụ nữ kiểu gen CB/cb là thể mang alen gây bệnh.

Alen *cb* khá hiếm gặp ở các quần thể người. Ở một số quần thể, tần số bệnh này ở nam khoảng 1/40. Còn ở nữ, vì phải có hai alen *cb* nên tần số càng thấp hơn, khoảng 1/1600, bằng bình phương tần số ở nam giới.

Người đàn ông mù màu phải được di truyền alen *cb* cùng với nhiễm sắc thể X từ mẹ. Nếu bà mẹ nhìn màu bình thường thì bà phải là dị hợp tử, tức là có kiểu gen *CB/cb*. Người đàn ông không thể truyền alen *cb* cho con trai vì con trai chỉ được di truyền nhiễm sắc thể Y từ bố. Ông ta chỉ có thể truyền alen *cb* cho các con gái. Bất kỳ người con gái nào nhận được alen *cb* từ bố và alen *CB* từ mẹ cũng là cá thể mang gen bệnh và sẽ truyền bệnh này cho một nửa số con trai. Con gái của một phụ nữ là thể mang bệnh với bố bị bệnh mù màu cũng có xác suất mắc bệnh này là 50% (*Hình 3.22*).Vì thế, kiểu hình ở bố sẽ không xuất hiện ở thế hệ tiếp theo, nhưng biểu hiện trở lại ở các cháu trai. Tất cả những người đàn ông bị bệnh, ngoại trừ những đột biến xuất hiện mới, đều được di truyền alen gây bệnh từ mẹ. Những dấu hiện trên đây là đặc điểm di truyền liên kết giới tính.

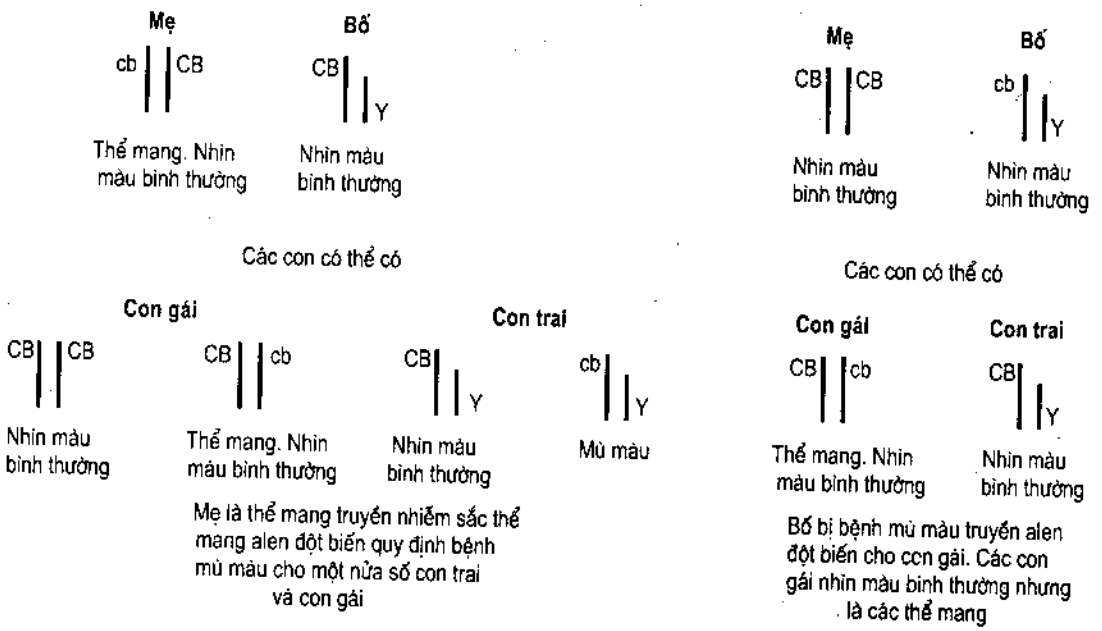
Hội chứng loạn dưỡng cơ Duchenne và bệnh máu khó đông cũng là những ví dụ khác về các bệnh di truyền liên kết giới tính ở người. Cụ thể, trong trường hợp bệnh máu khó đông, nghiên cứu phả hệ cho thấy Nữ hoàng Victoria có một alen đột biến gây bệnh máu khó đông

và truyền bệnh này cho một số con trai; đồng thời nhiều gia đình hoàng gia khác ở Châu Âu là thông gia kết hôn với các con gái bà cũng được truyền alen này. Vì không có bằng chứng chứng tỏ bệnh này xuất hiện từ các thế hệ trước của Nữ hoàng, nên người ta cho rằng Nữ hoàng Victoria đã nhận một alen đột biến xuất hiện trong giai đoạn phát sinh tế bào sinh dục của một trong hai bố, mẹ của bà.

Sự di truyền liên kết giới tính trở nên phức tạp hơn do có sự bất hoạt nhiễm sắc thể giới tính X ở giới cái ở các loài động vật có vú. Điều đó có nghĩa là: về lý thuyết, có khoảng một nửa số tế bào của cá thể cái mang alen bệnh sẽ biểu hiện bệnh, còn một nửa thì không. Ví dụ về hiện tượng này có thể thấy trong trường hợp *bệnh thiếu năng enzym gluco-6-phosphate* hoặc *hội chứng Lech-Nyhan* (khi người ta tiến hành phân tích từng tế bào). Tuy vậy, hiện tượng này thường không đủ tác động đến kiểu hình chung của cả cá thể mang gen.

Sự di truyền liên kết giới tính không được thể hiện đối với các gen nằm trên **vùng giả nhiễm sắc thể thường**, đó là vùng của nhiễm sắc thể X có thể kết cặp với nhiễm sắc thể Y trong quá trình giảm phân. Ở người, vùng này được tìm thấy ở đỉnh cánh ngắn của nhiễm sắc thể X. Vì có sự tương đồng giữa hai nhiễm sắc thể X và Y ở vùng này, nên hiện tượng trao đổi chéo có thể xảy ra giữa hai nhiễm sắc thể X và Y. Các gen ở vùng này có kiểu di truyền giống như các gen trên nhiễm sắc thể thường.

Bất kỳ gen nào nằm trên nhiễm sắc thể Y cũng được di truyền trực tiếp từ bố sang con trai. Tuy vậy, có một số ví dụ còn gây tranh cãi về kiểu di truyền này, như các tính trạng "người nhím" và chân có màng. Một gen đột biến quy định lông trên vành tai ngoài biểu hiện kiểu **di truyền thẳng**. Gần đây, các nghiên cứu cho thấy, gen xác định giới tính SRY và một gen mã hóa kháng nguyên hòa hợp tế bào chất phụ là H-Y nằm trên nhiễm sắc thể Y, được di truyền trực tiếp từ bố sang con trai.



Hình 3.22. **Sự di truyền bệnh mù màu. NST X được minh họa bằng đường vẽ dài, NST Y được biểu diễn bằng đường vẽ ngắn**

Các tính trạng bị giới hạn hoặc bị ảnh hưởng bởi giới tính

Các thuật ngữ này được dùng để chỉ các kiểu hình do cùng một kiểu gen quy định nhưng bị biến đổi phụ thuộc giới tính. Hai thuật ngữ này dễ gây nhầm lẫn nên cần được phân biệt rõ ràng.

Các alen của một gen bị giới hạn bởi giới tính chỉ biểu hiện kiểu hình ở một giới (hoặc đực hoặc cái). Ví dụ về kiểu tính trạng này là alen đột biến *BRCA1* quy định tính nhạy cảm với bệnh ung thư vú. Alen này là alen trội và chỉ biểu hiện ở giới cái. Trái lại, một gen quy định tính nhạy cảm với bệnh ung thư vú thứ hai là *BRCA2* có khả năng gây bệnh ung thư vú ở cả giới đực và cái, hay nói cách khác gen này không bị giới hạn bởi giới tính.

Sự khác biệt giữa các gen bị giới hạn bởi giới tính và bị ảnh hưởng bởi giới tính khó phát hiện, song có ý nghĩa quan trọng. Trong trường hợp thứ nhất ở ví dụ trên, kiểu hình chỉ giới hạn ở một giới. Trong khi đó, ở trường hợp thứ hai, kiểu hình được biểu hiện ở cả hai giới, nhưng phổ biến hơn ở giới này so với giới kia. Một ví dụ về kiểu di truyền này là chứng hói đầu. Chứng này do một gen gồm hai alen là *B* và *b* kiểm soát. Cá thể đồng hợp tử *BB* bị hói đầu, còn *bb* thì không. Dị hợp tử *Bb* có kiểu hình khác biệt ở nam và nữ. Ở nam giới, alen *B* biểu hiện trội và người có kiểu gen dị hợp tử bị hói đầu. Nhưng ở nữ giới, nó lại biểu hiện lặn và người phụ nữ *Bb* không bị hói (*Bảng 3.12*). Biểu hiện trội - lặn của alen này rõ ràng bị ảnh hưởng bởi sự cân bằng hormon giới tính trong cơ thể. Trong trường hợp này, đáng chú ý là gen liên quan đến chứng hói đầu cũng liên quan với bệnh buồng trứng có vách ở người phụ nữ.

Bảng 3.12. **Sự biểu hiện của các kiểu gen hói đầu ở nam và nữ**

Kiểu gen	Kiểu hình ở nữ	Kiểu hình ở nam
BB	Hói	Hói
Bb	Không	Hói
bb	Không	Không

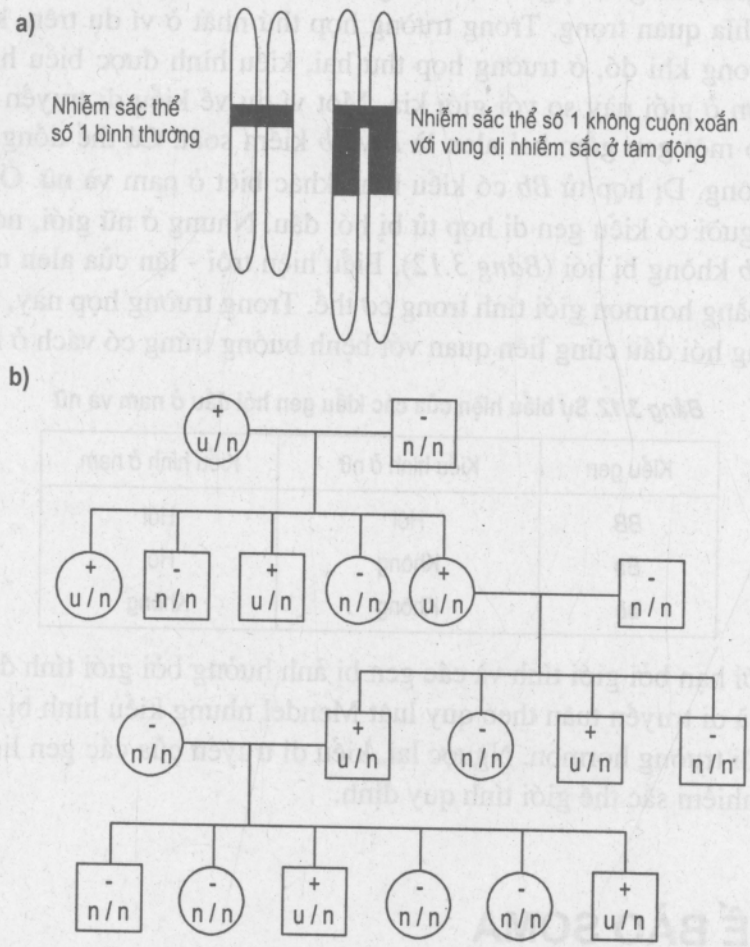
Các gen bị giới hạn bởi giới tính và các gen bị ảnh hưởng bởi giới tính đều là các gen nằm trên NST thường và di truyền tuân theo quy luật Mendel nhưng kiểu hình bị biến đổi có thể do sự biến đổi của môi trường hormon. Ngược lại, kiểu di truyền của các gen liên kết giới tính do các gen nằm trên nhiễm sắc thể giới tính quy định.

III.K LAI TẾ BÀO SOMA

Nhóm liên kết và nhiễm sắc thể

Các phương pháp lập bản đồ gen bằng phép lai phân tích chỉ cho phép xác định trình tự gen theo trật tự thẳng hàng trong các nhóm liên kết, chứ không cho phép xác định gen đó thuộc nhiễm sắc thể nào hoặc thuộc vùng nào của một nhiễm sắc thể. Để thực hiện được điều này, phải kết hợp phương pháp lập bản đồ gen với các phương pháp tế bào học. Trong một số trường hợp, có thể sử dụng tính đa hình di truyền tế bào xuất hiện trong các quần thể tự nhiên. Các chuyển đoạn NST làm thay đổi nhóm liên kết của các gen cũng có thể được dùng vì mục đích nghiên cứu tương tự. Đáng chú ý là sự thay đổi hình thái nhiễm sắc

thể có tương quan với việc xuất hiện hay không xuất hiện một kiểu hình được quy định bởi các alen khác nhau của cùng một gen. Một ví dụ về lập bản đồ gen người thực hiện vào năm 1969 được minh hoạ trên *Hình 3.23*. Hình thái nhiễm sắc thể số 1 thay đổi ở các quần thể người khác nhau do sự biến đổi vùng dị nhiễm sắc ở tâm động. Nhiễm sắc thể có tâm động dài được coi là "không cuộn xoắn", và đáng chú ý là trong một số gia đình người Mỹ gồm nhiều thế hệ sống chung, kiểu nhiễm sắc thể này chỉ xuất hiện ở những người dương tính với nhóm máu Duffy. Gen này có hai alen, một alen dương tính và alen kia âm tính đối với nhóm máu này. Các nghiên cứu trên những gia đình không có quan hệ họ hàng cho thấy: sự phân ly của gen này và nhiễm sắc thể không cuộn xoắn khẳng định sự liên kết của nhóm máu Duffy với nhiễm sắc thể số 1 và vị trí của gen gần vị trí tâm động.



Hình 3.23. **(a) Sơ đồ nhiễm sắc thể số 1 không cuộn xoắn và nhiễm sắc thể bình thường ở người. Nhiễm sắc thể không cuộn xoắn dài hơn do dư thừa vùng dị nhiễm sắc ở tâm động. (b) Sự phân ly của nhóm máu Duffy và nhiễm sắc thể số 1 không cuộn xoắn trong các thành viên một gia đình giả thiết. Cụ tổ dị hợp tử về cả hai tính trạng. Nhóm máu Duffy luôn phân ly cùng nhiễm sắc thể số 1 không cuộn xoắn và không tái tổ hợp. (+) Người dương tính về nhóm máu Duffy; (–) Người âm tính về nhóm máu Duffy; (u) Người có nhiễm sắc thể số 1 không cuộn xoắn; (n) Người có nhiễm sắc thể số 1 bình thường.**

Rõ ràng cách tiếp cận nêu trên có yếu tố may rủi và không phù hợp để tiến hành lập bản đồ gen trên nhiễm sắc thể một cách hệ thống, đặc biệt đối với loài người vốn có số lượng nhiễm sắc thể lớn, tần số sai hình nhiễm sắc thể ở những các cá thể hữu thụ thấp.

Một kỹ thuật hiệu quả hơn đã được sử dụng để xác định vị trí gen trên nhiễm sắc thể xuất phát từ quan sát của Weiss và Ephrussi (1967) là sử dụng các tế bào động vật nuôi cấy được dung hợp với nhau.

Nuôi cấy mô và sự dung hợp tế bào

Các tế bào động, thực vật có thể sinh trưởng trên môi trường nuôi cấy về cơ bản giống như các tế bào vi sinh vật thông thường. Đối với các tế bào động vật, cần tách mô vô trùng và đưa vào môi trường có chứa các amino acid, vitamin, nguồn năng lượng (thường là glucose) và bổ sung huyết thanh động vật. Tế bào người nuôi cấy có vòng đời ngắn và giới hạn cao nhất là khoảng 50 lần phân bào. Điều quan trọng là các tế bào này phải duy trì được số lượng lưỡng bội nhiễm sắc thể của chúng. Các tế bào của các loài gặm nhấm được nuôi cấy có thể thay đổi đặc tính trong điều kiện *invitro* và tiếp tục sinh trưởng không xác định trong môi trường nuôi cấy.

Sự dung hợp tế bào là quá trình có xảy ra trong tự nhiên, như sự kết hợp của tế bào trứng với tinh trùng trong thụ tinh, hay như sự kết hợp các nguyên bào cơ thành các ống cơ. Nó cũng có thể là kết quả của một quá trình lây nhiễm bởi virus, như trong trường hợp bệnh sởi hoặc quai bị. Những thí nghiệm đầu tiên về dung hợp tế bào nuôi cấy là sử dụng virus bất hoạt, rồi sau đó được thay thế bằng các tác nhân hóa học như polyethylen glycon (PEG). Trong điều kiện nuôi cấy *invitro*, sự dung hợp tế bào có thể được thực hiện giữa các tế bào cùng loài hay giữa các loài khác nhau, trong đó thậm chí các tế bào thực vật và động vật cũng có thể dung hợp với nhau. Sản phẩm đầu tiên của sự dung hợp là tế bào dị nhân (tế bào có chứa nhiều hơn một nhân có nguồn gốc di truyền khác nhau). Hầu hết các tế bào dị nhân không sinh trưởng được nhưng cũng có một tỷ lệ nhất định các tế bào lai sống sót. Có các phương pháp chọn lọc khác nhau đã được phát triển để có thể chọn ra các dòng tế bào lai trong hỗn hợp với các tế bào không dung hợp.

Bảng 3.13. **Nhiễm sắc thể bị mất ở các tế bào lai**

Tế bào lai giữa hai loài	Nhiễm sắc thể bị mất là của loài
Người - chuột nhà	Người
Người - chuột đồng	Người
Chuột cống - chuột nhà	Ổn định
Chuột đồng - chuột nhà	Chuột nhà
Khỉ - chuột đồng	Khỉ
Lợn - chuột đồng	Lợn
Cừu - chuột đồng	Cừu

Weiss và Ephrussi nhận thấy rằng, các tế bào lai cùng loài thường có số lượng nhiễm sắc thể ổn định, nhưng tế bào lai soma chuột - người có xu hướng không ổn định về đặc điểm này và số lượng nhiễm sắc thể có chiều hướng giảm đi khi tế bào lai phân chia. Sự giảm số lượng nhiễm sắc thể thường xảy ra đối với nhiễm sắc thể người. Ban đầu, nhiễm sắc thể người mất đi nhanh chóng nhưng khi số lượng nhiễm sắc thể người trong tế bào tiến dần tới không (0) thì tốc độ mất đi của nhiễm sắc thể chậm dần và các nhiễm sắc thể còn lại được duy trì khá ổn định. Hiện tượng tương tự cũng được thấy ở tế bào lai người - chuột đồng Trung Quốc, nhưng

ở đây nhiễm sắc thể người mất đi với tốc độ nhanh hơn nhiều. Cho dù cơ chế kiểm soát hiện tượng này còn chưa rõ, nhưng hiện tượng mất một bộ nhiễm sắc thể trong các tế bào lai soma khác loài là hiện tượng phổ biến. *Bảng 3.13* cho thấy bộ nhiễm sắc thể bị mất ở các tế bào lai giữa một số loài.

Sắp xếp gen trên các nhiễm sắc thể

Vì ở các tế bào lai soma giữa tế bào người và các loài gặm nhấm, các nhiễm sắc thể người có xu hướng bị mất. Nhưng, nhờ hiện tượng mất nhiễm sắc thể người là khác nhau trong các dòng tế bào lai khác nhau, nên các dòng tế bào lai này có thể được sử dụng để phân tích từng nhiễm sắc thể của người. Sau giai đoạn các nhiễm sắc thể người mất đi nhanh chóng, mỗi tế bào của từng dòng tế bào lai soma sẽ chỉ còn duy trì một số ít nhiễm sắc thể người. Nếu tách từng tế bào và nuôi cấy thì sẽ thu được các dòng tế bào chứa các nhiễm sắc thể người khác nhau. Các nhiễm sắc thể người có thể được xác định bằng phương pháp phân tích tế bào để lập bảng liệt kê các dòng tế bào lai. Trong đó, mỗi dòng chỉ mang một nhiễm sắc thể người duy nhất. Các dòng như vậy được minh hoạ trên *Bảng 3.14*.

Bảng 3.14. **Các dòng từ 1 - 11 là các dòng tế bào lai soma chuột - người, mỗi dòng có một bộ nhiễm sắc thể người duy nhất. Sự có mặt của nhiễm sắc thể người cụ thể ở các dòng tế bào lai được ký hiệu 'p', còn vắng mặt – 'a'. Sự có mặt của gen insulin người xác định bằng kỹ thuật Southern blotting được ký hiệu bằng 'i', vắng mặt – 'n'.**

Nhiễm sắc thể	Dòng 1	Dòng 2	Dòng 3	Dòng 4	Dòng 5	Dòng 6	Dòng 7	Dòng 8	Dòng 9	Dòng 10	Dòng 11
1	p	a	a	a	p	a	a	a	a	a	a
2	a	a	a	a	a	a	a	a	p	a	a
3	a	a	a	a	a	a	p	a	a	a	p
4	a	a	a	p	a	p	a	a	a	p	a
5	a	a	a	p	a	a	a	a	a	a	a
6	p	a	p	a	a	a	a	a	a	p	a
7	a	a	p	a	a	p	a	a	a	a	a
8	a	p	a	p	p	a	a	a	p	a	a
9	a	a	a	a	a	a	a	p	a	a	a
10	a	a	a	a	a	a	p	a	p	a	p
11	p	p	a	a	p	a	a	p	p	a	a
12	a	a	a	a	a	a	a	a	a	p	a
13	a	a	p	a	p	a	a	a	a	a	a
14	a	a	a	p	a	p	a	a	a	a	a
15	p	a	a	a	a	a	a	a	a	a	a
16	a	a	a	a	a	a	a	a	p	p	a
17	p	p	a	a	a	a	a	a	a	a	a
18	a	a	p	a	a	a	a	p	a	a	a
19	a	a	p	a	a	a	a	a	a	a	a
20	p	a	a	a	a	a	p	a	a	a	a
21	a	a	a	a	p	a	p	a	a	a	a
22	a	a	a	a	a	p	a	a	a	a	a
X	a	p	p	a	a	a	a	a	a	a	a
Insulin	i	i	n	n	i	n	n	i	i	n	n

Bảng này có thể được sử dụng để sắp xếp các gen người vào các nhiễm sắc thể bằng cách phân tích có xuất hiện hay không các gen cụ thể liên quan với các nhiễm sắc thể người xuất hiện ở mỗi dòng tế bào lai. Việc xác định gen insulin của người trên nhiễm sắc thể 11 được trình bày trong ví dụ ở *Bảng 3.14*. Sự có mặt hay vắng mặt gen insulin cũng có thể được xác định bằng phương pháp thẩm tách Southern với việc sử dụng gen insulin của người đã được tách dòng làm mẫu dò. Gần đây, kỹ thuật PCR được sử dụng để nhận biết sự có mặt của nhiều gen khác nhau ở người. Như chúng ta thấy trong ví dụ trên, gen insulin có mặt ở các dòng 1, 2, 5, 8, 9 và vắng mặt ở các dòng 3, 4, 6, 7, 10 và 11. Chỉ có nhiễm sắc thể 11 cho thấy mối tương đồng giữa sự có mặt và vắng mặt nhiễm sắc thể và sản phẩm của gen này. Hiện tượng đó được gọi là **sự phân ly hòa hợp** và là bằng chứng khẳng định gen insulin nằm trên nhiễm sắc thể 11. Khi một gen và nhiễm sắc thể không phân ly cùng nhau thì đó là **sự phân ly không hoà hợp**. Nếu hai gen phân ly hoà hợp thì có nghĩa là chúng nằm trên cùng một nhiễm sắc thể, chúng liên kết với nhau. Đáng chú ý là các gen liên kết có thể nằm xa nhau trên nhiễm sắc thể và không nhất thiết luôn biểu hiện sự di truyền liên kết với nhau.

Khi một gen được xếp vào một nhiễm sắc thể thì phương pháp phân tích tái tổ hợp sẽ cho thấy nó là một phần của nhóm các gen liên kết đã được xác định trên cùng nhiễm sắc thể đó. Kỹ thuật này có ảnh hưởng to lớn đến nghiên cứu di truyền học người trong giai đoạn 1970 - 1990. Phương pháp nghiên cứu tương tự cũng đã được sử dụng để nghiên cứu các gen của chuột trong các dòng tế bào lai soma [chuột × chuột đồng] mà ở đó các NST chuột có xu hướng bị mất đi.

Tuy vậy, đây chỉ là một phần của vấn đề. Cho dù nó là một bước quan trọng để sắp xếp các gen vào các nhiễm sắc thể nhưng không cho thông tin chính xác về vị trí xác định của gen trên nhiễm sắc thể. Cần tiếp tục cải tiến kỹ thuật để thúc đẩy việc lập bản đồ gen.

Xây dựng bản đồ gen chi tiết

Sự phân vùng các gen trên nhiễm sắc thể có thể thực hiện được bằng việc sử dụng một loạt các kỹ thuật phân tích các dòng tế bào lai soma khác loài chỉ mang một đoạn nhiễm sắc thể người.

Cách đơn giản nhất để thực hiện được mục đích nghiên cứu đó là tạo ra các tế bào lai soma với tế bào người thu được từ các cá thể mang một đột biến chuyển đoạn NST đã biết. Chẳng hạn như, nếu các tế bào người dùng để lai mang chuyển đoạn cân bằng giữa nhiễm sắc thể số 4 và nhiễm sắc thể X (trong đó cánh ngắn của nhiễm sắc thể X được chuyển sang nhiễm sắc thể số 4) thì việc phân tích các dòng tế bào lai đối với các gen trên nhiễm sắc thể X sẽ cho các tỷ lệ phân ly khác nhau. Các gen trên cánh ngắn của nhiễm sắc thể X sẽ phân ly hòa hợp với các gen trên nhiễm sắc thể số 4, trong khi các gen trên cánh dài sẽ được duy trì hoặc bị mất đi cùng với nhiễm sắc thể X bị đứt. Điều đó cho phép xác định gen thuộc cánh nào của nhiễm sắc thể X. Nhờ sử dụng nhiều chuyển đoạn khác nhau mà nhiều gen đã được sắp xếp vào các phân vùng trên các nhiễm sắc thể. Các điểm đứt gãy chuyển đoạn đã được dùng để chia nhiễm sắc thể 21 thành 19 phân vùng khác nhau.

Một phương pháp khác có thể dùng để tạo ra các đoạn nhiễm sắc thể người ở các tế bào lai là trực tiếp đưa vào tế bào lai các đoạn nhiễm sắc thể được lấy ở kỳ giữa. Các NST bị tổn thương trong quá trình xử lý như vậy và chỉ có những đoạn nhỏ được kết hợp vào tế bào lai. Cũng có thể tạo ra các tế bào lai chỉ mang một nhiễm sắc thể người duy nhất và phân lập các

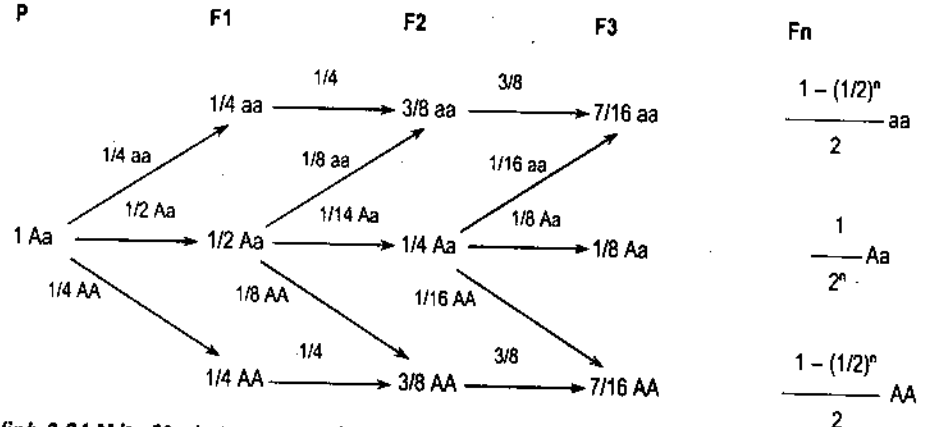
dòng trong đó các đoạn NST này đã bị mất. Các gen bị mất phải nằm trên vùng nhiễm sắc thể bị mất đi đó.

Kỹ thuật dung hợp tế bào soma đã cung cấp một phương tiện hữu hiệu để lập bản đồ gen ở người và một số loài sinh vật khác nhau. Tuy nhiên, nên chú ý là kỹ thuật này chỉ giới hạn với các gen có được các mẫu dò phân tử là các dòng cADN, hoặc có trình tự ADN đã biết để có thể thực hiện phản ứng PCR đặc hiệu. Các gen quy định các hội chứng di truyền mà các khiếm khuyết hoá sinh cơ bản chưa xác định được thì không lập được bản đồ theo cách này.

III.L NỘI PHỐI

Nội phối

Nghiên cứu nội phối là nghiên cứu hiệu ứng của sự giao phối giữa các cá thể có quan hệ họ hàng gần. Trường hợp tự phối của các sinh vật lưỡng tính là trường hợp nội phối cực đoan nhất. Hiện tượng này khá phổ biến ở thực vật. Nội phối làm tăng tính đồng hợp tử (*Hình 3.24*). Trong trường hợp tự phối, tỷ lệ dị hợp tử giảm xuống một nửa sau mỗi thế hệ vì ở các thể dị hợp tử, một nửa số tổ hợp giao tử là đồng hợp tử (cụ thể là, $Aa \times Aa$ sẽ sinh ra $1AA : 2Aa : 1aa$). Mối nguy hiểm ở đây là các alen lặn có hại có mặt với tần số thấp trong quần thể sẽ trở thành đồng hợp tử ở đời con. Hiện tượng đó làm giảm sức sống và được gọi là **sự ức chế nội phối**. Đối với thực vật tự thụ phấn (như lúa mỳ, đậu), hiện tượng này không nguy hiểm vì chọn lọc tự nhiên gần như đã loại bỏ hết các alen lặn có hại. Thực tế, hiện tượng nội phối còn góp phần giúp quá trình chọn lọc của người nông dân vì các cây trồng đời con đồng hợp tử sẽ giống hệt bố, mẹ về mặt di truyền. Sản lượng và đặc tính của cây trồng có thể dự đoán trước và ổn định từ năm này sang năm khác, đồng thời mang các đặc điểm nông sinh học giống nhau (như kích thước, thời điểm chín, chất lượng). Đối với các sinh vật giao phối xa, sự ức chế nội phối có thể gây ra hậu quả nghiêm trọng. Việc lai giữa hai dòng nội phối để tạo ra đời con F1 gồm toàn các thể dị hợp tử là cần thiết. Sự khôi phục lại các đặc tính tốt được gọi là sức sống con lai nhưng nó không được truyền cho toàn bộ các cá thể ở F2 vì có hiện tượng phân ly kiểu hình ở F2. Ở cây ngô, các dòng thuần chủng thường rất yếu và người ta đã dùng bốn dòng thuần để tạo ra các dòng lai F1 rồi lại lai tiếp với nhau để tạo ra hạt lai kép làm thương phẩm.



Hình 3.24. **Mức độ gia tăng tính đồng hợp tử qua các thế hệ liên tục tự thụ phấn (sự nội phối cực đoan)**

Hiện tượng nội phối sinh ra mức đồng hợp tử quá ngưỡng. Nhiều locut gen sẽ ở trạng thái đồng hợp tử chỉ đơn giản vì ít có biến dị di truyền. Trong các quần thể giao phối ngẫu nhiên, một alen có tần số q sẽ tồn tại ở trạng thái đồng hợp tử với tần số là q^2. Hầu hết các đột biến đặc biệt chỉ xuất hiện một lần. Nếu có được một cây phả hệ đầy đủ, chúng ta sẽ nhận thấy những người mang alen đột biến thường đều xuất phát từ một tổ tiên chung. Hiện tượng giao phối giữa các cá thể có quan hệ họ hàng tương đối gần làm cho tần số đồng hợp tử cao hơn so với giá trị mong đợi trên cơ sở quy luật Hardy-Weinberg. Hiện tượng này được đặc trưng bằng hệ số nội phối, đó là xác suất tự hợp tử - một trường hợp đặc biệt của thể đồng hợp tử, khi cả hai alen của một locut đều được di truyền từ một tổ tiên chung.

Hiệu ứng đối với tần số alen

Trong các quần thể nội phối, các alen lặn có hại trở thành đồng hợp tử và chịu tác động của chọn lọc nhanh hơn khi chúng tồn tại trong một quần thể lớn ngoại phối. Các thể đồng hợp tử về các alen lặn có hại thường nhanh chóng bị chọn lọc đào thải. Vì vậy mà các alen lặn có hại thường có tần số thấp trong các quần thể nội phối. Tất cả các đột biến có hại xuất hiện trong các quần thể nội phối sẽ nhanh chóng chuyển về trạng thái đồng hợp tử và bị đào thải sớm hơn do quá trình chọn lọc.

Hệ số đồng huyết

Hôn phối cận huyết là hôn phối giữa các cá thể có quan hệ họ hàng. Hệ số đồng huyết là xác suất để hai cá thể có quan hệ họ hàng cùng mang các alen giống hệt nhau do cùng được di truyền từ một tổ tiên chung. Mức độ đồng huyết của họ bằng hệ số nội phối ở đời con. Tuy vậy, điều đó chỉ sử dụng để xem xét các alen đặc biệt. Điều này có ý nghĩa trong y học vì nếu biết được một tổ tiên (ví dụ ông hoặc bà) có alen lặn có hại thì có thể tiên đoán được xác suất để các con cháu sinh con là đồng hợp tử về alen đó.

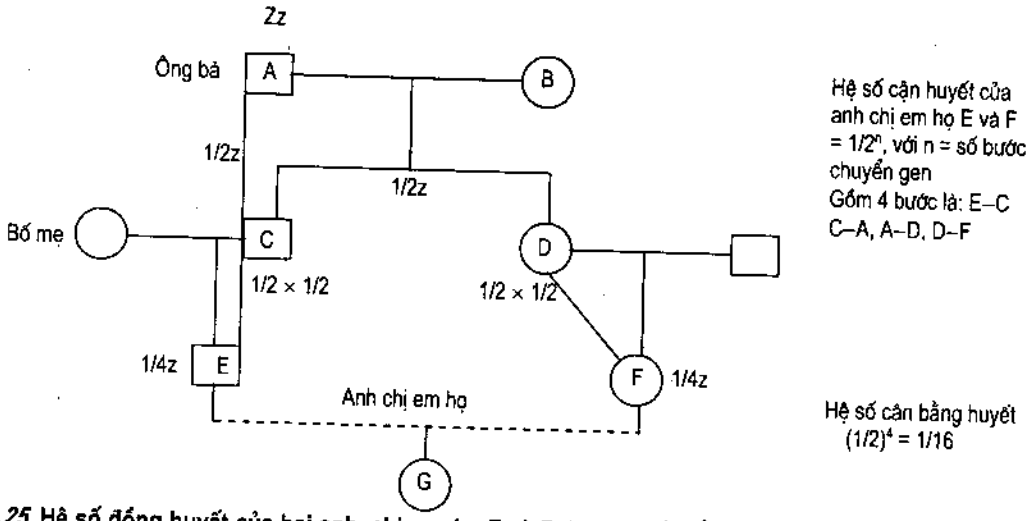
Hệ số cận huyết được tính như minh họa dưới đây. Mỗi giao tử nhận một nửa số alen từ bố, mẹ nên con sẽ có một nửa số alen được di truyền từ bố và một nửa từ mẹ; vậy xác suất để hai anh, chị, em nhận được cùng một bản sao của một alen là bao nhiêu? *Hình 3.25* cho thấy cách tính hệ số cận huyết. Xác suất để con trai C nhận được alen *z* từ bố là 1/2 (hay 0,5) và con gái D cũng có xác suất như vậy để nhận được alen z, nên xác suất để cả hai anh, chị, em cùng nhận được alen *z* là $(1/2)^2 = 1/4$. Đó chính là hệ số cận huyết của họ (chú ý rằng xác suất để họ cùng được di truyền alen *Z* cũng là 1/4 nên xác suất để cả hai cùng mang các bản sao của mỗi alen thực sự là 1/2). Tiếp tục theo phả hệ, C có được alen *z* với xác suất 1/2, và sẽ truyền alen *z* cho cháu trai E với xác suất 1/2 nên E sẽ có xác suất để có alen *z* là $1/2 \times 1/2 = 1/4$. Mức độ đồng huyết của cháu trai E với cô D là $1/4 \times 1/2 = 1/8$ (= 0,125). (Bác C và cháu F cũng vậy).

Tương tự, xác suất để D có alen *z* và truyền nó cho cháu F bằng $1/2 \times 1/2 = 1/4$. Do đó, hệ số đồng huyết của hai cháu E và F là: $1/4 \times 1/4 = 1/16$ hoặc 0,0625.

Công thức tính là:

$$\text{Hệ số đồng huyết} = \frac{1}{2^n}$$

Với n là số bước truyền gen (số bước tạo giao tử giữa họ), ví dụ: trên *Hình 3.25*, các bước đó là E-C, C-A, A-D và D-F (= 4 bước). (Có thể viết là E-C-A-D-F).



Hình 3.25. **Hệ số đồng huyết của hai anh, chị, em họ E và F là xác suất để cả hai cùng được di truyền một bản sao của cùng một alen (giả sử z) từ một tổ tiên chung cụ thể (ông nội A) bằng 1/16, bằng với hệ số nội phối của con họ (G). Hình tròn: nữ; hình vuông: nam**

Hệ số nội phối

Hệ số nội phối (F) là xác suất để một cá thể có cả hai alen của một gen giống hệt nhau do được di truyền từ cùng một alen ở tổ tiên gần nhất. Hệ số nội phối có thể được tính theo đường di truyền và theo xác suất để alen đó được di truyền từ cả hai bố và mẹ. *Hình 3.26a* cho thấy đường truyền đến con trong cuộc hôn nhân cận huyết anh, chị, em ruột. Xét alen z_i ở ông nội A. Xác suất để nó được truyền vào giao tử tạo nên người bố C là 1/2, và nó cũng được truyền cho đứa con I (sib) của người bố này với xác suất 1/2, nên xác suất người con nhận được alen z_i từ bố là $(1/2)^2$. Mặt khác, alen này cũng có cùng xác suất truyền cho người mẹ D. Nếu người con nhận được alen z_i từ cả hai bố, mẹ thì sẽ là đồng hợp tử. Xác suất cho sự kiện đó là:

$$(1/2)^2_{(bố)} \times (1/2)^2_{(mẹ)} = (1/2)^4 = 1/16$$

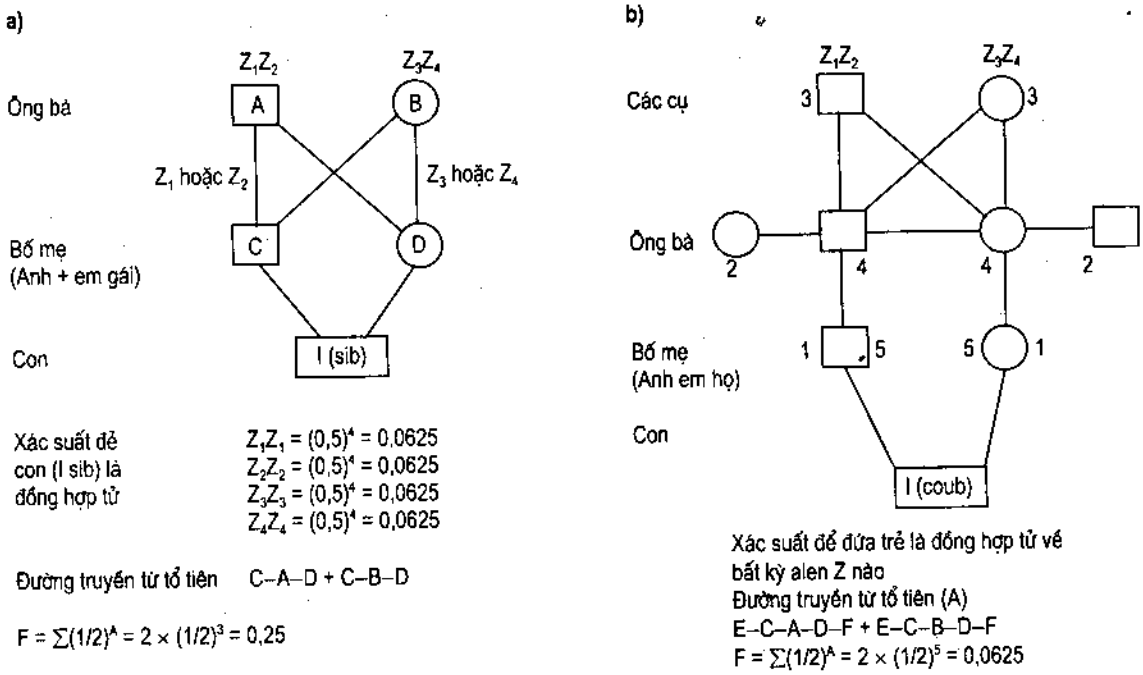
Tuy nhiên, vì có bốn alen khác nhau ở ông bà (z_1, z_2, z_3 và z_4) nên đứa trẻ I (sib) có khả năng như nhau để là đồng hợp tử về bất kỳ alen nào trong số đó. Vì vậy, xác suất để nó là đồng hợp tử về bất kỳ alen nào ở locut này (chính là hệ số nội phối của nó) là $F = 4 \times (1/2)^4 = 1/4$. (Hệ số này có trị số giống như hệ số đồng huyết giữa bố, mẹ C và D). Theo cách đó để xét đứa con gái của cuộc hôn nhân anh, chị, em họ ($I_{(cous)}$, *Hình 3.26b*), chúng ta thấy cách truyền gen khác từ tổ tiên nên xác suất để cô con gái này nhận được một alen cụ thể giảm còn $(1/2)^2$ và hệ số nội phối là $4 \times (1/2)^6 = (1/2)^4 = 1/16$.

Công thức để tính F là:

– Hệ số nội phối = $F = \sum(1/2)^A$

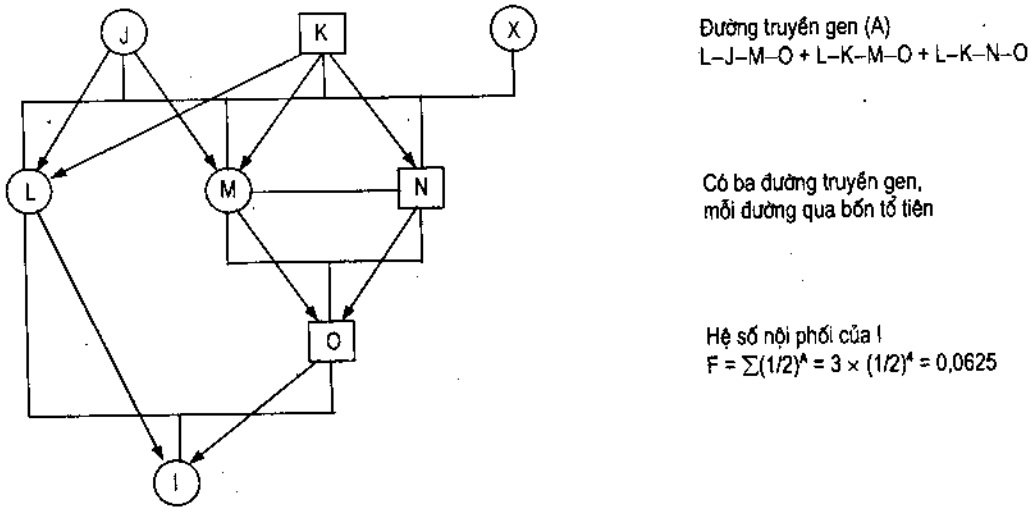
– Với A là số tổ tiên mà mỗi đường truyền gen đi qua một tổ tiên chung và giá trị của các đường truyền được cộng lại. Ví dụ, đối với cuộc hôn nhân anh, chị, em ruột ($I_{(sib)}$, *Hình 3.26a*), các đường truyền là C-A-D và C-B-D, nên A = 3 cho cả hai đường và $F_{(sib)} = (1/2)^3 + (1/2)^3 = 1/4$. Trong trường hợp hôn nhân giữa anh, chị, em họ ($I_{(cous)}$, *Hình 3.26b*), cả hai đường truyền đều đi qua 5 tổ tiên: qua tổ tiên A là E-C-A-D-F và qua B là E-C-B-D-F. Vậy A = 5 cho mỗi đường truyền và:

$$F_{(cous)} = (1/2)^5 + (1/2)^5 = 1/16$$



Hình 3.26. **(a) Hệ số nội phối F của con I của hai bố, mẹ C và D là anh, chị, em ruột. (b) Hệ số nội phối F của cuộc hôn nhân giữa anh, chị, em họ. Chữ số bình thường từ trái qua phải là chuỗi các tổ tiên thông qua A, chữ số nhỏ, nghiêng, từ phải qua trái là chuỗi các tổ tiên thông qua B. (Hình tròn: nữ; hình vuông: nam)**

Trong các chương trình chọn giống, có thể có những trường hợp phức tạp hơn, ví dụ như trên *Hình 3.27*. Có ba đường truyền, mỗi đường truyền qua bốn tổ tiên là L-J-M-O, L-K-M-O và L-K-N-O, vậy $F_{(I)} = 3(1/2)^4 = 3/16$.



Hình 3.27. **Một kiểu di truyền phức tạp hơn từ chương trình chọn giống. Các cá thể J và K sinh ra L và M; K cũng sinh ra N; Hai anh, chị, em cùng cha khác mẹ (N x M) sinh ra O; O giao phối với L sinh ra I có hệ số nội phối 3/16. Cá thể X không đóng góp gì vào hệ số nội phối. (Mũi tên chỉ đường truyền alen. Hình tròn: cái; hình vuông: đực)**

Tầm quan trọng đối với y học

Ở trạng thái cân bằng Hardy-Weinberg, tần số alen lặn, hiếm gặp là q và tần số đồng hợp tử lặn là q^2. Phần lớn các bệnh di truyền là do các alen lặn, hiếm gặp, ở trạng thái đồng hợp tử gây ra. Nếu một alen có tần số $q = 0{,}0001$ (hay 1/10 000) thì $q^2 = 1 \times 10^{-8}$ (hay 1/100.000.000). Tuy nhiên, đối với đứa trẻ có bố, mẹ là anh, chị, em họ, thì tần số đó là $pqF = 0{,}0001 \times 0{,}9999 \times 1/16 \approx 6{,}25 \times 10^{-6}$ (hay 1/160.000). Như vậy, tần số đồng hợp tử tăng 625 lần so với quần thể bình thường. Mức gia tăng đó bằng pF/q. Phần lớn các bệnh di truyền hiếm gặp tìm thấy ở người là do giao phối cận huyết gây nên. Các nghiên cứu cũng cho thấy rằng những cuộc hôn nhân giữa anh, chị, em họ có khả năng sảy thai hoặc tử vong khi sơ sinh lớn hơn gấp từ 2 đến 10 lần so với các cuộc hôn nhân bình thường.

III.M XÁC SUẤT

Xác suất

Trong nghiên cứu di truyền học, người ta thường phải tính tần số và số lượng các kiểu gen mong đợi. Chẳng hạn như việc phải ước đoán tần số kiểu hình và kiểu gen trong các phép lai theo các nguyên tắc di truyền học Mendel, hoặc dự đoán tần số của các kiểu gen riêng biệt trong một quần thể trên cơ sở định luật Hardy - Weinberg. Các dự đoán về tần số kiểu gen như vậy được xác nhận bằng các dữ liệu thực nghiệm sẽ giúp khẳng định hoặc đánh giá các giả thiết.

Chúng ta biết rằng, khi tung một đồng xu rơi xuống đất, xác suất để một trong hai mặt ngửa lên trên trong mỗi lần tung là 0,5 (tức 50%), nhưng chúng ta không thể khẳng định mặt nào chắc chắn sẽ xuất hiện trong mỗi lần tung bởi vì mỗi lần tung đồng xu là một sự kiện xảy ra độc lập và kết quả xảy ra là ngẫu nhiên. Tương tự như vậy khi thực hiện một phép lai, chúng ta có thể biết tỷ lệ số hoa có màu hồng nhưng không thể khẳng định chính xác cây hoa nào có màu hồng hoặc màu trắng. Để biết kết quả, chúng ta chỉ có cách là thực hiện phép lai. **Lý thuyết xác suất** cho phép chúng ta tính được tỷ lệ (tần số) các kiểu gen và kiểu hình mong đợi.

Xác suất (P) của một sự kiện sẽ xảy ra là tỷ số giữa số lần sự kiện đó xảy ra (a) chia cho tổng số lần sự kiện đó có thể xảy ra trong thực tế (n)

$$P = a/n$$

P có thể được xác định trực tiếp hoặc ngoại suy từ bản chất của sự kiện. Chúng ta có thể giả sử khi tung đồng xu, xác suất để một mặt nhất định ngửa lên trên là 0,5. Tương tự như vậy, xác suất để một mặt nhất định của hạt xúc sắc xuất hiện trong mỗi lần gieo xúc sắc là 1/6 (bởi vì khả năng xuất hiện của một trong sáu mặt 1, 2, 3, 4, 5 và 6) là như nhau. **Tổng số xác suất của tất cả các sự kiện có thể xảy ra phải bằng 1**, bởi vì lúc nào cũng **luôn có một** trong các sự kiện xảy ra (ví dụ: khi gieo xúc sắc, chắc chắn sẽ có một trong sáu mặt ngửa lên trên). Nếu một sự kiện luôn luôn xảy ra, thì xác suất của nó bằng 1, còn nếu nó không bao giờ có khả năng xảy ra thì xác suất bằng 0.

Nguyên tắc cộng xác suất

Nguyên tắc cộng xác suất được áp dụng đối với các sự kiện có ảnh hưởng tương quan qua lại lẫn nhau. Khi gieo xúc sắc, mặt xuất hiện có thể là mặt 5, cũng có thể là mặt 6, nhưng không bao giờ xuất hiện cả hai cùng lúc. Vậy, xác suất để xuất hiện hoặc mặt 5, hoặc mặt 6 là:

$$P_{(5 \text{ hoặc } 6)} = 1/6 + 1/6 = 1/3$$

Một ví dụ về xác suất trong di truyền là phép lai thu được tỷ lệ di truyền Mendel về màu hoa là 1/4 đỏ : 2/4 hồng và 1/4 trắng. Như vậy, xác suất để một bông hoa bất kỳ có màu đỏ hoặc màu hồng sẽ là 1/4 + 2/4 = 3/4.

Nguyên tắc nhân xác suất

Nguyên tắc nhân xác suất được áp dụng đối với các sự kiện xảy ra riêng lẻ hoặc các sự kiện xảy ra theo một trật tự nhất định. Chẳng hạn, xác suất để nhận được hai mặt 6 cùng một lúc khi gieo hai hạt xúc sắc khác nhau là bao nhiêu? Việc gieo hạt xúc sắc này là độc lập với việc gieo hạt xúc sắc kia. Xác suất xuất hiện mặt 6 khi gieo hạt xúc sắc thứ nhất là 1/6, tương tự như vậy đối với hạt xúc sắc thứ hai. Vì vậy, xác suất xuất hiện đồng thời hai mặt 6 là:

$$P_{(6 \text{ và } 6)} = 1/6 \times 1/6 = 1/36$$

Một ví dụ trong di truyền học là trường hợp kết hôn giữa hai người có kiểu gen dị hợp tử về một bệnh di truyền do alen lặn quy định là phenylketo niệu (PKU). Nếu đôi vợ chồng này có 3 đứa trẻ, thì xác suất ngẫu nhiên để ba đứa trẻ đều biểu hiện bệnh (đồng hợp tử lặn) là bao nhiêu? Vì xác suất mỗi đứa trẻ nhận được kiểu gen đồng hợp tử lặn trong mỗi lần sinh là 1/4, nên xác suất để cả ba đứa trẻ đều có kiểu gen này là:

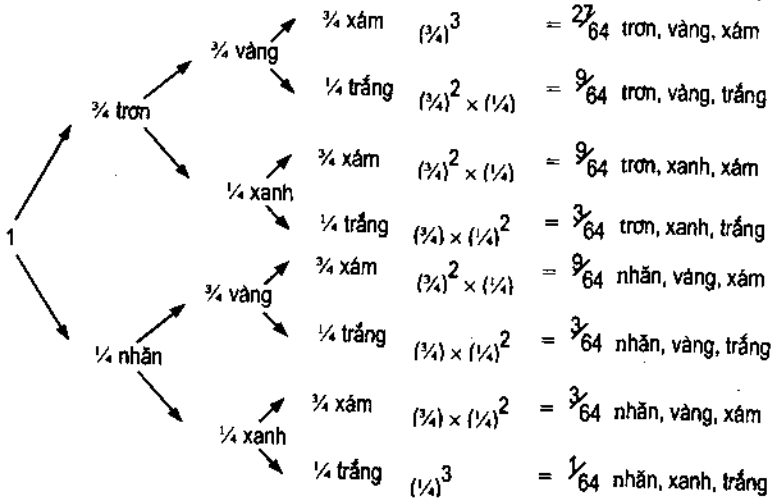
$$P_{(3 \text{ đứa trẻ đều bị PKU})} = 1/4 \times 1/4 \times 1/4 = 1/64$$

Tính xác suất

Các sự kiện độc lập không bị ảnh hưởng bởi kết quả của các sự kiện xảy ra độc lập trước đó. Chẳng hạn trong gia đình có hai vợ chồng mang alen PKU vừa nêu ở trên, xác suất để cả ba đứa con đều biểu hiện bệnh là 1/64 chỉ thực sự đúng khi đôi vợ chồng này chưa sinh con. Nếu đôi vợ chồng này đã có 2 đứa con bị bệnh, thì xác suất đứa con thứ ba cũng bị bệnh không hề thay đổi mà vẫn là 1/4 do sự kiện sinh đứa con thứ ba là một sự kiện độc lập. Tương tự như vậy, giả sử chúng ta tung đồng xu và nhận được cùng một mặt ngửa lên trên trong 10 lần liên tiếp, thì trong lần tung thứ 11, xác suất lại xuất hiện mặt đó vẫn là 0,5.

Nguyên tắc cộng và nhân xác suất thường được áp dụng đồng thời. Chẳng hạn xác suất để một cặp vợ chồng có một con trai và một con gái sẽ bằng tổng xác suất có con trai đầu lòng và con gái thứ hai (1/2 × 1/2 = 1/4) với xác suất có con gái đầu lòng và con trai thứ hai (1/2 × 1/2 = 1/4), hay nói cách khác là (1/4 + 1/4 =) 1/2. Sự phân ly của mỗi locut gen (nằm trên các nhiễm sắc thể khác nhau) là một sự kiện xảy ra độc lập (theo định luật II của Mendel). Cách thức đơn giản nhất để tính tần số phối hợp của các sự kiện xảy ra độc lập (kiểu gen hay kiểu hình) là thông qua nguyên tắc tính xác suất. Việc lập bảng Punnett đôi khi phức tạp và có thể nhầm lẫn. Chẳng hạn, giả sử như chúng ta thực hiện phép lai ba yếu tố ở cây đậu Hà Lan như thí nghiệm của Mendel. Nếu chúng ta tiến hành lai giữa một dòng thuần mang ba tính trạng trội hạt tròn, vàng, có lớp vỏ hạt màu xám (RRYYGG) với một dòng mang ba tính trạng lặn hạt nhăn, xanh, có lớp vỏ màu trắng (rryygg), con lai F1 (RrYyRg) có khả năng tạo ra 8 kiểu giao tử khác nhau. Nếu lập bảng Punnett trong trường hợp này, chúng ta sẽ có 64 ô. Lúc đó việc áp dụng nguyên tắc tính xác suất để trực tiếp xác định tần số kiểu gen thế hệ F2 sẽ trở nên đơn giản hơn. Chẳng hạn, xác suất một cá thể F2 có kiểu hình hạt tròn, xanh và vỏ hạt màu xám có thể được tính một cách đơn giản bởi nguyên tắc xác suất. Xét tổng số, có 3/4 số

cá thể có hạt tròn (tròn là trội), 1/4 có hạt xanh (xanh là lặn) và 3/4 có vỏ hạt màu xám (vỏ hạt xám là trội). Như vậy, xác suất để một cá thể có hạt tròn, màu xanh và vỏ hạt màu xám sẽ là 3/4 tròn × 1/4 xanh × 3/4 vỏ xám = 9/64. Tổng xác suất của tất cả các kiểu hình có thể xảy ra là 64/64 = 1 (xem *Hình 3.28*). Nguyên tắc tương tự cũng có thể áp dụng đối với kiểu gen.



Hình 3.28. **Xác suất tần số kiểu hình ở F2 trong một phép lai ba tính theo nguyên tắc di truyền học Mendel được tính theo nguyên tắc nhân xác suất**

Ví dụ: tần số kiểu gen RryyGg ở thế hệ F2 là bao nhiêu? Ta có tần số Rr là 1/2, tần số của yy là 1/4 và của Gg là 1/2, vì vậy tần số của RryyGg sẽ là 1/2 × 1/4 × 1/2 = 1/16 (hoặc 4/64 theo *Bảng Punnett*).

Phép hoán vị

Phép hoán vị là cách sắp xếp thứ tự các yếu tố khác đi nhưng kết quả cuối cùng không thay đổi. Khi chúng ta gieo hai viên xúc sắc để thu được hai mặt 6 chỉ có một phương án duy nhất là mặt 6 của cả hai viên xúc sắc đều phải ngửa lên trên. Tuy vậy, nếu chúng ta muốn có một mặt 3 và một mặt 6 thì có hai cách. Cách thứ nhất là viên xúc sắc thứ nhất mặt 3, viên thứ hai mặt 6. Cách thứ hai là viên xúc sắc thứ nhất mặt 6, viên thứ hai mặt 3. Đó chính là hai cách hoán vị:

$$P_{(3 \text{ và } 6)} = 1/6 \times 1/6 \times 2 = 1/18$$

Xác suất để một cặp vợ chồng có con trai và con gái là như nhau (P = 1/2), vì vậy xác suất để có cả hai con là trai sẽ là $(1/2)^2 = 1/4$. Tương tự như vậy, xác suất để có cả hai con là gái cũng sẽ là 1/4. Nhưng xác suất để có một con trai và một con gái sẽ là 1/2 × 1/2 × 2 = 1/2. Bởi vì, hai người con có thể được sinh ra theo thứ tự trai - gái hoặc gái - trai. Đây chính là hai cách hoán vị (hay còn gọi là cách **tổ hợp**). Trong trường hợp hai bố, mẹ dị hợp tử về gen PKU có ba người con, thì xác suất một trong ba đứa trẻ của họ bị bệnh (hai đứa con còn lại bình thường) sẽ là bao nhiêu?

Thực tế, đứa trẻ bị bệnh có thể là con đầu, con thứ hai hoặc con út, như vậy có ba cách hoán vị khác nhau. Xác suất để một đứa con của họ bị bệnh (A) và hai đứa con còn lại bình thường (N) sẽ là:

$$P_{(1A+2N)} = P\,(\text{A-N-N}) + P\,(\text{N-A-N}) + P\,(\text{N-N-A})$$
$$= (1/4 \times 3/4 \times 3/4) + (3/4 \times 1/4 \times 3/4) + (3/4 \times 3/4 \times 1/4)$$
$$= 3[(3/4)^2 \times 1/4] = 27/64.$$

Số khả năng hoán vị (hay tổ hợp) được tính dựa trên **hàm nhị thức mở rộng** $(p + q)^n$. Trong đó, p là xác suất của một sự kiện này (ví dụ: xác suất đứa trẻ bình thường là 3/4) còn p là xác suất của sự kiện còn lại (xác suất đứa trẻ bị bệnh PKU là 1/4) và n là số sự kiện có thể diễn ra (trong trường hợp này là tổng số đứa trẻ). Chú ý rằng $(p + q) = 1$. Nếu có n đứa trẻ, xác suất để có s đứa trẻ bình thường và t đứa trẻ bị bệnh sẽ là:

$$P = (n!/s!t!)p^s q^t$$

Số khả năng hoán vị được tính trong biểu thức $n!/s!t!$ và được gọi là **hệ số hoán vị**, còn $p^s q^t$ là xác suất xảy ra của các sự kiện theo một thứ tự nhất định. Chú ý rằng $(p + q) = 1$ và $(s + t) = n$. Ký hiệu ! là phép tính **giai thừa** (đọc là: giai thừa của n), nghĩa là phép nhân liên tục từ số nguyên n với các số nguyên giảm dần liên tục cho đến 1 (ví dụ: $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$). $0! = 1$, tương tự như vậy $n^0 = 1$. Nếu chúng ta tính lại xác suất trong trường hợp một đứa trẻ bị bệnh ($t = 1$) và hai đứa trẻ bình thường ($s = 2$), chúng ta có:

$$P = (n!/s!t!)p^s q^t$$
$$P = (3!/2!\,1!) \times (3/4)^2 \times (1/4)^1$$
$$= [(3 \times 2 \times 1)/\{(2 \times 1) \times 1\}] \times 9/16 \times 1/4$$
$$= 3 \times 9/64 = 27/64$$

Như vậy, kết quả trùng khớp với kết quả được tính ở trên, nhưng với việc áp dụng công thức này, việc tính toán thực hiện đơn giản hơn, chẳng hạn khi tính xác suất trong gia đình có 7 người con có 3 người con bị bệnh ($n = 7$, $s = 4$, $t = 3$).

Tam giác Pascal (*Hình 3.29*) cho biết hệ số hoán vị dựa trên hàm nhị thức mở rộng có thể sử dụng để áp dụng vào công thức tính xác suất nêu trên. Mỗi một hàng đều luôn bắt đầu và kết thúc bởi hệ số 1, và mỗi hệ số sẽ là tổng của hai hệ số liền kề với nó ở hàng trên. Trong ví dụ gồm ba đứa trẻ được nêu ở trên ($n = 3$), chúng ta sử dụng hàng thứ tư ($= n + 1$) trong tam giác Pascal (bởi vì chúng ta có bốn khả năng khác nhau tương ứng với số đứa trẻ có khả năng bị bệnh là 0, 1, 2 và 3). Hàng thứ tư trong tam giác Pascal cho thấy có một cách để số đứa trẻ bị bệnh là 0, có ba cách để số đứa trẻ bị bệnh là 1, có ba cách để số đứa trẻ bị bệnh là 2, và có một cách để số đứa trẻ bị bệnh là 3. Phương trình tính xác suất có thể được mở rộng từ hàm nhị thức mở rộng thành **hàm đa thức mở rộng**. Nếu xác suất để màu hoa đỏ, hồng và trắng tương ứng là p, q và r, thì xác suất để thu được s hoa đỏ, t hoa hồng và u hoa trắng là:

$$P = (n!/s!t!u!)p^s q^t r^u$$

Trong đó, tổng số hoa = $n = s + t + u$. Trong thực tế, công thức này có thể mở rộng cho số lượng sự kiện bất kỳ.

Tam giác Pascal

Hàng										Tổng số
1					1					1
2				1	1					2
3			1	2	1					4
4		1	3	3	1					8
5	1	4	6	4	1					16
6	1	5	10	10	5	1				32
7	1	6	15	20	15	6	1			64
8	1	7	21	35	35	21	7	1		128
9	1	8	28	56	70	56	28	8	1	256

Hình 3.29. **Tam giác Pascal cho biết Hệ số nhị thức mở rộng**

III.N PHƯƠNG PHÁP χ^2 VÀ KIỂM TRA GIẢ THIẾT

Kiểm tra giả thiết

Trong nghiên cứu di truyền học, một việc làm cần thiết là phải kiểm tra xem một tập hợp các số liệu thực nghiệm có trùng khớp với các số liệu lý thuyết (giả thiết) hay không. Chẳng hạn, trong phép lai đơn tính, tỷ lệ phân ly kiểu hình theo định luật phân ly của Mendel là 3 : 1. Giả sử chúng ta tiến hành lai giữa một cây hoa màu đỏ với một cây hoa màu trắng, tiến hành tự thụ phấn ở F1 (tất cả các cây có hoa màu đỏ) và ở F2 thu được 63 cây hoa đỏ : 37 cây hoa trắng. Câu hỏi đặt ra là: liệu tỷ lệ 63 : 37 có trùng khớp với tỷ lệ 3 : 1 hay không? Phép thử "khi bình phương" (χ^2) là một phép toán thống kê cho chúng ta biết xác suất để thu được kết quả thực nghiệm trùng với kết quả lý thuyết nếu như giả thiết trùng khớp là đúng. Nhưng phép thử này không thể cho biết giả thiết trùng khớp có thực sự đúng hay không. Giả thiết trùng khớp coi số liệu thực nghiệm là trùng với số liệu lý thuyết (số liệu giả thiết). Nghĩa là, trong thí nghiệm nêu trên, trước khi thực hiện phép thử kiểm tra giả thiết, chúng ta phải có giả thiết trùng khớp là tỷ lệ 63 : 37 là khớp với 3 : 1. ả

Phép thử χ^2

Giá trị χ^2 được tính bằng công thức:

$$\chi^2 = \text{tổng}\left(\frac{(\text{số liệu quan sát - số liệu giả thiết})^2}{\text{số liệu giả thiết}}\right)$$

hoặc được viết tắt là: $\chi^2 = \sum\left(\frac{(O-E)^2}{E}\right)$

χ^2 luôn được tính bằng số liệu quan sát được, chứ không được tính thông qua tỷ lệ phần trăm, tần số hay các tỷ lệ chia khác.

Việc tính χ^2 để so sánh tỷ lệ thực nghiệm 63 : 37 như được nêu trong ví dụ trên được trình bày trên *Bảng 3.15*. Số liệu lý thuyết được tính bằng việc nhân tổng số với tần số lý thuyết (ví dụ: [63 + 37] x 1/4 cây hoa màu trắng). Sự khác biệt giữa số liệu lý thuyết và thực nghiệm được lấy giá trị bình phương để loại bỏ số âm và chia cho giá trị lý thuyết để bù lại phép tính bình phương và để có thể so sánh với các số liệu tiêu chuẩn (số liệu trong bảng χ^2). Phép thử này được áp dụng cho mọi lớp số liệu (ví dụ là các nhóm cây hoa đỏ và hoa trắng như trong thí nghiệm trên) và được tính tổng để có số liệu χ^2 tổng quát. Chú ý rằng giá trị của χ^2 tăng lên khi sự khác biệt giữa thực nghiệm và giả thiết là lớn, vì vậy giá trị χ^2 lớn cho phép chúng ta loại bỏ giả thiết trùng khớp. Hay nói cách khác, số liệu thực nghiệm không phù hợp giả thiết khi giá trị χ^2 lớn. Kết quả thực nghiệm hoàn toàn trùng khớp với giả thiết khi giá trị χ^2 bằng 0. Theo kết quả tính toán ở *Bảng 3.15*, kết quả tính χ^2 là 7,68. Điều này có nghĩa là gì? Chúng ta sẽ so sánh giá trị này với giá trị trên *Bảng* χ^2. Nhưng trước hết, chúng ta cần phải biết bậc tự do.

Bậc tự do

Bậc tự do là số các lớp số liệu trừ đi 1. Bậc tự do cho chúng ta biết có bao nhiêu lớp số liệu có thể thay đổi tùy ý. Trong trường hợp có hai lớp số liệu như thí nghiệm trên (tức là cây hoa màu đỏ và màu trắng), thì chúng ta thấy tất cả các cây không phải hoa màu đỏ sẽ có hoa màu trắng. Khi chúng ta đếm số

cây có hoa đỏ, thì số cây có hoa trắng là cố định (tính trên tổng số), vì vậy bậc tự do là 1. Nếu chúng ta có ba lớp số liệu, chẳng hạn hoa đỏ, hoa hồng và hoa trắng, chúng ta sẽ có số bậc tự do bằng 2; nghĩa là khi đếm hai lớp số liệu, thì lớp số liệu thứ ba là cố định. Khi chúng ta làm phép thử đối với định luật Hardy - Weinberg với một locut và 2 alen (có 3 kiểu gen), chúng ta chỉ có 1 bậc tự do. Bởi vì số liệu lý thuyết được tính từ tần số thực nghiệm vốn phụ thuộc vào nhau (trong khi tỷ lệ dự đoán 3 : 1 là độc lập với tỷ số thực nghiệm). Chỉ có một lớp số liệu độc lập tương ứng với tần số của một alen (ví dụ p), còn tần số của alen còn lại và các kiểu gen bị cố định bởi p ($p + q = 1$, và dị hợp tử là $2pq$).

Bảng 3.15. **Ví dụ về cách tính χ^2**

Giả thiết trùng khớp (Tỷ lệ mong đợi)	Thực nghiệm		
3 : 1	Số liệu thực nghiệm (O) (tổng số: 100)	63 đỏ	37 trắng
	Số liệu mong đợi (số liệu giả thiết, E)	100 x 3/4 = 75	100 x 1/4 = 25
	O - E	- 12	12
	$(O - E)^2$	144	144
	$(O - E)^2/E$	144/75 = 1,92	144/25 = 5,76
$\chi^2 = 1,92 + 5,76 = 7,68$			
Bậc tự do = 2 - 1 = 1			
Từ *Bảng 3.16* ta có xác suất (p) tương ứng với bậc tự do 1 có $0,01 > p > 0,005$			
9 : 7	Số liệu thực nghiệm (O) (tổng số: 100)	63 đỏ	37 trắng
	Số liệu mong đợi (số liệu giả thiết, E)	100 x 9/16 = 56,25	100 x 7/16 = 43,75
	O - E	6,75	- 6,75
	$(O - E)^2$	45,5625	45,5625
	$(O - E)^2/E$	45,5625/56,25 = 0,81	45,5625/43,75 = 1,04
$\chi^2 = 0,81 + 1,04 = 1,85$			
Bậc tự do = 2 - 1 = 1			
Từ *Bảng 3.16* ta có xác suất (p) tương ứng với bậc tự do 1 có $0,20 > p > 0,05$			

Số bậc tự do có ảnh hưởng đến giá trị χ^2 bởi vì giá trị được cộng thêm vào mỗi khi một lớp số liệu mới được bổ sung. Vì vậy, giá trị χ^2 sẽ tăng lên khi số lớp số liệu tăng lên. Lúc đó, chúng ta sẽ phải tìm giá trị χ^2 lý thuyết từ một dòng khác trong bảng để phù hợp với số bậc tự do.

Sử dụng bảng xác suất χ^2

Bây giờ, chúng ta xem xét đến mức ý nghĩa thống kê trong kết quả phân tích của chúng ta. Quan sát *Bảng 3.16* và theo dòng 1 (tương ứng với bậc tự do 1) để tìm giá trị χ^2 lớn hơn và nhỏ hơn gần nhất so với giá trị χ^2 thực nghiệm vừa tính được (7,68), chúng ta thấy giá trị nhỏ hơn là 6,6349 tương ứng với xác suất 0,01 (tức là 1%, hay 1/100) và giá trị lớn hơn là 7,87944 tương ứng với xác suất là 0,005 (tức là 0,5% hay 5/1000). Giá trị χ^2 thực nghiệm nằm

ở giữa hai giá trị này. Điều đó có nghĩa là: nếu chúng ta thực hiện thí nghiệm lặp đi lặp lại 100 lần, chúng ta không thể có được sự trùng hợp số liệu thực nghiệm như vậy một lần nữa. Điều đó cho thấy số liệu thực nghiệm không trùng khớp với tỷ lệ 3 : 1.

Mức ý nghĩa thống kê

Mức ý nghĩa thống kê thông thường được chấp nhận là 5% (0,05). Nếu sự khác biệt giữa số liệu thực nghiệm và lý thuyết lớn đến mức mà xác suất xuất hiện ngẫu nhiên nhỏ hơn 5%, thì có thể loại bỏ giả thiết trùng khớp bởi lúc đó chúng ta có thể nói số liệu thực nghiệm khác biệt rõ rệt **có ý nghĩa thống kê** so với số liệu lý thuyết (số liệu mong đợi). Trong trường hợp nếu giá trị χ^2 thực nghiệm cao hơn giá trị χ^2 lý thuyết ở P = 0,05 (1/20) thì chúng ta sẽ đúng khi loại bỏ giả thiết trùng khớp trong 19/20 lần thực nghiệm. "Xác suất" không có nghĩa là "khẳng định" một sự kiện xảy ra, mà chúng chỉ cung cấp thông tin cho chúng ta về khả năng để sự kiện đó có thể xảy ra như mong đợi (giả thiết). Trong trường hợp trên, nếu chúng ta loại bỏ giả thiết trùng khớp giữa hai tỷ lệ 63 : 37 và 3 : 1, thì giả thiết theo định luật Mendel (3 : 1) vẫn có thể đúng nhưng tỷ lệ 63 : 37 sẽ xuất hiện ngẫu nhiên ít hơn 1 lần, nếu thí nghiệm được lặp lại cứ 100 lần.

Bảng 3.16. **Bảng giá trị phân phối χ^2 với các xác suất P tương ứng**

Bậc tự do \ P	0,995	0,975	0,050	0,025	0,010	0,005	0,001
1	392704×10^{-10}	982069×10^{-9}	3,84146	5,02389	6,63490	7,87944	10,828
2	0,0100251	0,0506356	5,99146	7,37776	9,21034	10,5966	13,816
3	0,0717218	0,215795	7,81473	9,34840	11,3449	12,8382	16,266
4	0,206989	0,484419	9,48773	11,1433	13,2767	14,8603	18,467
5	0,411742	0,831212	11,0705	12,8325	15,0863	16,7496	20,515
6	0,675727	1,23734	12,5916	14,4494	16,8119	18,5476	22,458
7	0,989256	1,68987	14,0671	16,0128	18,4735	20,2777	24,322
8	1,34441	2,17973	15,5073	17,5345	20,0902	21,9550	26,125
9	1,73493	2,70039	16,9190	19,0228	21,6660	23,5894	27,877

Giá trị ứng dụng của phép thử χ^2 là để khẳng định khả năng tin cậy của một kết luận. Chẳng hạn, nếu xác suất thu được một giá trị χ^2 nhất định nào đó là thấp hơn 0,01 thì nó thường được trích dẫn ở mức 1% hơn là ở mức 5%. Mức 1% có nghĩa là việc loại bỏ giả thuyết trùng khớp chỉ có thể sai với tần số là 1 lần trong 100 lần thí nghiệm.

Có thể có một giả thiết khác là tỷ lệ 63 : 37 có thể liên quan đến hai gen khác nhau ở hai giống hoa màu đỏ và màu trắng. Nếu những gen này mã hóa cho các enzym hoạt động trên cùng con đường sinh hóa thì chúng cần cùng lúc để có thể tạo ra sắc tố đỏ. Nếu giả thiết trùng khớp này là đúng thì tỷ lệ lý thuyết sẽ là 9 : 7 (xem mục III.B). Giá trị χ^2 tính theo giả thiết này cũng đã được trình bày trên *Bảng 3.15* và có giá trị là 1,85. *Bảng 3.16* cho chúng ta thấy giá trị χ^2 sẽ thấp hơn 1,64 ở P = 0,20 (tức là một phần năm số thí nghiệm) và lớn hơn 3,84 ở P = 0,05 (tức là 5% hay một phần hai mươi số thí nghiệm). Kết quả thu được của chúng ta là

xác suất nằm giữa hai mức xác suất này. Do sự sai khác thực nghiệm có thể xảy ra với tần số tương tự hoặc lớn hơn với tần số 1 lần trong 20 lần thí nghiệm, nên chúng ta có thể chấp nhận giả thiết trùng khớp và số liệu thực nghiệm là phù hợp với tỷ lệ 9 : 7 lý thuyết và giả thiết hai gen liên quan đến việc quy định tính trạng màu hoa đỏ có thể được chấp nhận. Có một khả năng nào đó là kết luận này sai, nhưng hiện tại chưa có giả thiết nào khác tỏ ra thuyết phục hơn, vì vậy ta có thể tạm thời chấp nhận.

Phép sửa sai số Yates

Phép sửa sai số Yates có thể áp dụng cho mọi lớp số liệu có số lượng dữ liệu nhỏ hơn 5. χ^2 là một hàm phân phối liên tục (nên chúng được biểu hiện bởi các giá trị thập phân), trong khi các số liệu thực nghiệm của chúng ta là các số nguyên (bởi chúng ta không thể trồng một nửa cây đậu). Bởi vậy, với số lượng dữ liệu thấp, theo lý thuyết chúng ta có thể có 3,5 cây hoa màu trắng, nhưng số liệu thực nghiệm gần nhất có thể thu được là 3 hoặc 4, như vậy sai số tối thiểu là 0,5. Phép sửa sai số Yates được áp dụng bằng cách giảm giá trị chênh lệch (số liệu thực nghiệm – số liệu lý thuyết) đi 0,5 (theo hướng tiến tới 0) nếu số lượng dữ liệu nhỏ hơn hoặc bằng 5. Nếu giá trị chênh lệch là -1,6 sẽ được hiệu chỉnh về -1,1; nếu giá trị chênh lệch là 2,7 thì được hiệu chỉnh về 2,2 trước khi lấy bình phương và áp dụng vào biểu thức tính χ^2.

Phép thử chính xác Fisher

Phép thử chính xác Fisher được áp dụng với những trường hợp có ít sự kiện hoặc khi một số hoặc tất cả các lớp sự kiện xảy ra với tần số thấp (hiếm gặp). Trong những trường hợp này, phép thử Fisher cho phép tính được xác suất gần với kết quả thực nghiệm hơn. Xác suất của một sự kiện ngoại lệ nào đó có thể tính được từ xác suất của mỗi sự kiện riêng lẻ. Xác suất tổng số của các sự kiện riêng lẻ được tính để đưa ra xác suất có độ lệch lớn hơn hoặc bằng độ lệch quan sát được ngẫu nhiên. Nếu xác suất này nhỏ hơn 0,05, chúng ta sẽ loại bỏ giả thiết trùng khớp. Chẳng hạn, nếu giả thiết trùng khớp là những đứa trẻ sinh ra là trai hay gái là như nhau (xác suất là 0,5 với mỗi trường hợp) thì xác suất để có hai đứa trẻ sinh ra liên tiếp là gái sẽ chính xác là $0,5^2 = 0,25$. Vì vậy, nếu hai đứa trẻ sinh ra liên tiếp là gái, thì điều đó không khác biệt đáng kể với tỷ lệ giới tính mong đợi là 1 : 1. Nhưng xác suất để nhận được 6 đứa con liên tiếp đều là gái sẽ là $0,5^6 = 0,015625$ thì việc một gia đình có 6 đứa con đều là gái (giả sử chúng ta không hề biết về sự có mặt của những gia đình có cả 6 người con đều là trai) thì chúng ta dễ dẫn đến kết luận loại bỏ giả thiết rằng xác suất sinh con trai hay gái là như nhau với mức ý nghĩa thống kê là 1,5625%.

Bảng 3.17. **Phép thử chính xác về kiểu hình độ dài lông ở chó con**

Độ dài lông (& số con)		Xác suất riêng lẻ x số khả năng hoán vị	Xác suất kết quả lai mỗi loại	Xác suất tích lũy
Dài (n)	Ngắn (n)			
4	0	1/4 x 1/4 x 1/4 x 1/4	0,00390625	0,00390625
3	1	1/4 x 1/4 x 1/4 x 1/4 x 4	0,046875	0,05078125
2	2	1/4 x 1/4 x 3/4 x 3/4 x 6	0,2109375	0,26171875
1	3	1/4 x 3/4 x 3/4 x 3/4 x 4	0,421875	0,68359375
0	4	3/4 x 3/4 x 3/4 x 3/4	0,31640625	1

Chúng ta có thể xem xét một ví dụ phức tạp hơn: một cặp chó lông ngắn được lai với nhau và đẻ ra 4 con chó con, có 3 con lông dài và một con lông ngắn. Như vậy, cặp chó này phải là dị hợp tử, và tính trạng lông ngắn là trội và giả thiết trùng khớp sẽ tương ứng tỷ lệ phân ly mong đợi ở thế hệ con phải là 3 lông ngắn : 1 lông dài. Vậy, xác suất để nhận được một trường hợp có kết quả ngược lại một cách ngẫu nhiên là bao nhiêu? Xác suất để xảy ra sự kiện với độ lệch lớn hơn hoặc bằng sự kiện như vậy là bao nhiêu? Trong những trường hợp này, số liệu thực nghiệm quá nhỏ để có thể áp dụng phép thử χ^2. Trường hợp xảy ra với độ lệch lớn nhất so với sự kiện mong chờ sẽ là cả bốn con chó con đều có lông dài, vì vậy chúng ta sẽ bắt đầu với trường hợp đó. *Bảng 3.17* cho chúng ta thấy xác suất riêng lẻ và xác suất tổng số của tất cả các khả năng phối hợp có thể xảy ra với các con chó con có lông dài hoặc lông ngắn trong trường hợp lứa đẻ có bốn con. Xác suất để một lứa đẻ có ba con chó con lông dài và một con lông ngắn, hoặc cả bốn con lông dài là 0,05078125, gần tương đương với tỷ lệ 1/20, vì vậy chúng ta không có lý do gì để loại bỏ giả thiết được nêu ở trên. Thậm chí nếu trường hợp tất cả các con chó con là lông ngắn (xác suất là 0,0039; tức là xấp xỉ 1/256), thì chúng ta thực sự nên lặp lại thí nghiệm vì số sự kiện xảy ra trong thí nghiệm (lứa đẻ có 4 con) không đủ lớn để có thể đưa ra kết luận tổng quát. Phép thử chính xác Fisher cũng được sử dụng để kiểm tra sự trùng hợp giữa kết quả thực nghiệm và lý thuyết khi phân tích di truyền quần thể theo định luật Hardy - Weinberg đối với các alen và kiểu gen xảy ra với tần số hiếm. Một ứng dụng quan trọng khác của phép thử này là ước tính mức độ nội phối và liên kết gen (sự liên kết của các alen từ các locut khác nhau) trong các quần thể người để kiểm chứng giả thiết về tần số các kiểu gen trong nghiên cứu hình sự. Một số kiểu gen chứa các alen có thể xuất hiện với tần số hiếm đến mức không phát hiện ra dù đã phân tích hàng ngàn cá thể. Có quá nhiều khả năng kết hợp để có thể tính toán một cách chính xác về xác suất xuất hiện của chúng. Phép thử chính xác lúc đó phải được mô phỏng bằng máy tính để có thể cùng lúc phân tích hàng ngàn cách thức tổ hợp của các alen với các tần số của chúng ghi nhận từ thực tế. Tần số các alen sẽ được tính toán xấp xỉ với tần số chính xác để các kiểu gen lạ có thể xảy ra dựa trên định luật Hardy - Weinberg.

Phần IV
DI TRUYỀN HỌC QUẦN THỂ VÀ TIẾN HÓA

Phần IV gồm các mục:

IV.A Giới thiệu

Giới thiệu chung
Chọn lọc tự nhiên
Sự cân bằng Hardy - Weinberg
Đa dạng di truyền
Thuyết Darwin mới
Sự tiến hóa nhiễm sắc thể
Loài và sự hình thành loài
Đa bội thể
Sự tiến hóa và quần thể

IV.B Tiến hóa nhờ chọn lọc tự nhiên

Tiến hóa nhờ chọn lọc tự nhiên
Những quan sát và kết luận của Darwin
Phương thức chọn lọc
Hình thái loài khởi sinh
Chọn lọc giới tính
ADN/gen ích kỷ
Điểm lại lịch sử

IV.C Trạng thái cân bằng Hardy - Weinberg

Mở đầu
Tính tần số alen
Tần số các kiểu gen ở trạng thái cân bằng
Tác nhân gây sai lệch trạng thái cân bằng
Đánh giá độ phù hợp
Ước tính tần số alen

IV.D Đa dạng di truyền

Mở đầu
Nguồn biến dị
Chọn lọc chống lại đột biến
Lạc dòng di truyền
Chọn lọc phụ thuộc tần số
Chọn lọc cân bằng - ưu thế dị hợp tử

IV.E Thuyết Darwin mới: chọn lọc tác động đến các alen

Mở đầu
Đơn vị chọn lọc
Mức độ phù hợp (hay Giá trị chọn lọc)
Sự nhiễm melanin công nghiệp
Chọn lọc chống lại các alen lặn
Chọn lọc chống lại các alen trội
Chọn lọc chống lại các alen cộng gộp
Chọn lọc ưu tiên dị hợp tử
Chọn lọc chống lại các thể dị hợp tử

IV.F Sự biến đổi nhiễm sắc thể trong quá trình tiến hóa

Kiểu nhân
Hiện tượng đa hình nhiễm sắc thể và tính hữu thụ
Các cơ chế tái cấu trúc nhiễm sắc thể
Mất đoạn
Lặp đoạn
Sự kết hợp tâm động và phân đôi tâm động
Chuyển đoạn
Đảo đoạn
Đảo đoạn không tâm
Đảo đoạn mang tâm
Sự biến đổi ở nhiễm sắc thể giới tính
Hiệu ứng tiến hóa

IV.G Loài và sự hình thành loài

Loài
Các cơ chế cách ly sinh sản
Sự hình thành loài
Sự tiếp xúc trở lại
Các mô hình phát sinh chủng loại

IV.H Đa bội thể

Giới thiệu
Tự đa bội
Dị đa bội
Sự kết hợp giữa các hệ gen
Các phức loài đa bội
Đặc tính của các dạng đa bội thể

IV.I Quá trình tiến hóa

Tiến hóa do sự phân ly di truyền
Các quần thể
Các loài địa lý
Đồng hồ phân tử
Cây chủng loại phát sinh
Sự tiến hóa của loài người

IV.A GIỚI THIỆU

Giới thiệu chung

Di truyền học quần thể và tiến hoá liên quan mật thiết với nhau. Di truyền học quần thể nghiên cứu về mặt số lượng tần số alen và tần số kiểu gen trong quần thể, trong khi đó tiến hoá là sự biến đổi các tần số đó qua thời gian. Các nhân tố làm biến đổi tần số alen cũng chính là các nhân tố gây nên tiến hoá. Chính vì thế mà các mục của phần này cũng có mối liên quan mật thiết, khó tách biệt.

Chọn lọc tự nhiên

Chọn lọc tự nhiên kiểm định lại các quan sát và kết luận của Darwin. Những quan sát và kết luận đó là cơ sở để tìm hiểu các nhân tố gây nên những biến đổi có tính thích nghi về mặt di truyền trong quần thể để đáp lại môi trường của chúng, và những biến đổi đó qua thời gian chính là sự tiến hoá. Bất kỳ môi trường sống nào cũng chỉ nuôi dưỡng được một quần thể có giới hạn của một loài; khả năng đó thường nhỏ hơn nhiều so với tiềm năng sinh sản của quần thể. Vì các cá thể cạnh tranh nhau để sinh sản, nên những cá thể truyền được cho con cái chúng vật chất di truyền của mình sẽ là những cá thể thành công nhất. Như vậy, sự tiến hoá diễn ra.

Trạng thái cân bằng Hardy -Weinberg

Hardy và Weinberg, độc lập với nhau, cùng thực hiện những tính toán cho thấy sự cân bằng về tần số các kiểu gen sẽ đạt được sau một thế hệ ngẫu phối, và tần số các kiểu gen có thể tính được xuất phát từ tần số alen. Tần số các kiểu gen của trạng thái cân bằng sẽ được duy trì từ thế hệ này sang thế hệ khác cho đến khi có một áp lực nào đó làm thay đổi tần số alen. Hardy và Weinberg đã bác bỏ nhận xét sai lầm rằng các tính trạng trội sẽ trở nên phổ biến hơn. Việc đưa vào sử dụng những phép tính số học đơn giản đã khởi đầu cho môn Di truyền học quần thể và cho phép chúng ta định lượng hiệu ứng tiến hoá ở những mức độ khác nhau đối với các thể đồng hợp tử và dị hợp tử, hoặc chính là đối với các alen trội và lặn. Trạng thái cân bằng là cơ sở của thuyết Darwin mới, đồng thời cũng là cơ sở để tìm hiểu sự phân bố các bệnh di truyền ở các quần thể người.

Đa dạng di truyền

Các cá thể của hầu hết các quần thể của các loài đều rất đa dạng về mặt di truyền. Nguồn biến dị này là do đột biến; nhưng tại sao lại có sự đa dạng lớn như vậy? Có hai cách giải thích: là do chọn lọc tự nhiên và sự xảy ra ngẫu nhiên. Chọn lọc tự nhiên liên tục thay thế các alen hoặc có thể là sự chọn lọc phụ thuộc vào tần số alen nhằm ưu tiên cho các alen hiếm để giữ chúng tồn tại trong quần thể. Cả hai cách giải thích đều đòi hỏi phải có một số lượng lớn cá thể bị chết ở mỗi thế hệ gây ra do chọn lọc tự nhiên. Tuy vậy, có quá nhiều biến dị đến mức chỉ bằng chọn lọc tự nhiên thì không giải thích nổi. Có lẽ, phần lớn biến dị không ảnh hưởng đến giá trị chọn lọc (hay mức độ thích nghi) của cá thể, vì thế chúng là trung tính về mặt chọn lọc và tần số các alen đó biến động hoàn toàn ngẫu nhiên. Cách giải thích đúng nhất có lẽ là có sự kết hợp nào đó giữa hai yếu tố, tuy vậy vẫn còn nhiều tranh cãi về tầm quan trọng của chúng.

Thuyết Darwin mới

Sự kết hợp giữa di truyền học quần thể và chọn lọc tự nhiên cho phép chúng ta đánh giá hiệu quả của chọn lọc đối với các kiểu gen và tiên đoán kết quả chọn lọc. Nó làm cho việc nghiên cứu tiến hoá

trở thành khoa học định lượng. Chúng ta có thể tính toán số thế hệ cần thiết để một kiểu chọn lọc cụ thể theo alen trội hoặc lặn dẫn đến sự thay thế alen này bằng alen kia. Chúng ta thấy rằng, các alen lặn ít bị tác động hơn vì chúng chủ yếu ẩn ở trạng thái dị hợp tử, và sự chọn lọc duy trì dị hợp tử sẽ bảo tồn cả hai alen trội và lặn, trong khi chọn lọc chống lại dị hợp tử sẽ loại bỏ các alen hiếm, thậm chí chúng thích nghi khi ở trạng thái đồng hợp tử. Điều đó dẫn chúng ta đến việc xem xét chọn lọc như sự tác động đến alen chứ không phải là tác động đến kiểu hình và xây dựng quan điểm về các "gen ích kỷ" đảm bảo cho sự di truyền của riêng nó chứ không đảm bảo cho giá trị chọn lọc của cá thể.

Sự tiến hóa nhiễm sắc thể

Các nhiễm sắc thể mang thông tin di truyền. Hình dạng và số lượng của chúng được gọi là **kiểu nhân** của loài, là đặc trưng cho mỗi loài, nhưng có thể thay đổi theo thời gian tiến hoá. Sự phát sinh giao tử đòi hỏi có sự phân chia chính xác của vật chất di truyền trong giảm phân; điều đó khó thực hiện được nếu các nhiễm sắc thể không tương đồng. Có những loài mà tính dị hợp tử về hình dạng nhiễm sắc thể là phổ biến; nhiều phần của nhiễm sắc thể bị đảo đoạn, hoặc chuyển đoạn sang nhiễm sắc thể khác. Một số hiệu ứng di truyền đặc biệt xuất hiện từ những hiện tượng đó. Tái tổ hợp (trao đổi chéo) trong nguyên phân ở các thể dị hợp tử có sự tái cấu trúc nhiễm sắc thể thường sinh ra các sản phẩm không có khả năng sống, vì thế mà hiện tượng tái tổ hợp bị cản trở ở những đoạn nhiễm sắc thể có tái cấu trúc và tất cả các alen ở vùng đó có thể cùng tiến hoá. Trong một số trường hợp khác, sự khác biệt nhiễm sắc thể gây nên tính bất thụ ở con lai và hoạt động như một tác nhân theo hướng hình thành loài mới. Việc so sánh nhiễm sắc thể giữa các loài có thể cho thấy sự biến đổi qua tiến hoá.

Loài và sự hình thành loài

Sự hình thành loài mới là sự kiện chủ yếu trong quá trình tiến hoá. Các cá thể của những loài khác nhau không giao phối với nhau và không thể cùng chia sẻ một vốn gen chung. Việc khống chế sự trao đổi vật chất di truyền là yêu cầu duy nhất quan trọng cho sự hình thành loài, và cuối cùng, nó phải được kiểm soát về mặt di truyền khi hai loài gặp gỡ nhau và có cơ hội để giao phối. Nếu hai loài bị cách biệt về mặt vật lý theo thời gian giao phối hoặc khoảng cách không gian thì chúng không thể trao đổi vật chất di truyền với nhau. Nếu xuất hiện những biến đổi trung tính hoặc có chọn lọc thì những biến đổi đó chỉ được duy trì trong quần thể. Cuối cùng, theo thời gian, sự khác biệt ngày càng lớn, tới mức sự giao phối giữa các thành viên của hai quần thể không thành công được; vậy là sự hình thành loài mới diễn ra.

Đa bội thể

Có những cá thể mang ba bộ nhiễm sắc thể hoặc nhiều hơn. Chúng là các thể đa bội, khác với các thể lưỡng bội chỉ có hai bộ nhiễm sắc thể. Những sinh vật có ba bộ nhiễm sắc thể là các thể tam bội, có bốn bộ nhiễm sắc thể là tứ bội và ngũ bội (5 bộ), lục bội (6 bộ), bát bội (8 bộ)... Các thể đa bội thường xuất hiện dưới dạng con lai giữa hai loài thân thuộc. Chúng mang một bộ nhiễm sắc thể từ mỗi bố, mẹ và không tương đồng với nhau. Các thể đa bội lẻ bất thụ vì chúng không thể phân tách chính xác các nhiễm sắc thể vào các giao tử trong giảm phân: nhiễm sắc thể lẻ không thể phân đôi được. Tuy nhiên, bộ nhiễm sắc thể đó có thể ngẫu nhiên được nhân đôi trong nguyên phân, khi đó mỗi nhiễm sắc thể có một cặp và con lai trở nên hữu thụ. Sự kiện đó là phổ biến ở thực vật. Có đến 47% thực vật có hoa là thể đa bội. Vì các giao tử của chúng

chỉ tạo ra hợp tử cân bằng gen khi kết hợp với giao tử của thể đa bội giống như nó, nên các cá thể đa bội khởi đầu có thể hình thành nên loài mới nhờ cơ chế cách ly chức năng giữa các loài.

Sự tiến hóa và quần thể

Việc so sánh các loài cho thấy những gì đã xảy ra trong quá khứ. Những loài có tổ tiên chung gần đây thường giống nhau hơn các loài phân ly từ xa xưa. Việc so sánh trình tự amino acid của những protein cụ thể cho thấy chúng biến đổi với những tốc độ khác nhau. Những protein tương tác với nhiều phân tử khác phải duy trì cấu trúc của nó để đảm bảo hoạt động chức năng, và do vậy biến đổi chậm. Các protein khác thuộc loại hoà tan được và chủ yếu tương tác với nước thì ít bền vững và biến đổi nhanh hơn. Việc kiểm tra ADN cho thấy những trình tự không mã hoá biến đổi nhanh nhất. Theo cách đó chúng ta có thể xây dựng đồng hồ phân tử, dựa trên tốc độ đột biến thực của các base được chọn lọc tự nhiên sàng lọc chống lại các đột biến có hại. Kết quả cho thấy rằng hầu hết các đột biến duy trì khả năng sống sót của sinh vật là các đột biến trung tính. Kết quả đó còn cho chúng ta tiên đoán khoảng cách tiến hoá giữa các loài bằng cách so sánh trình tự ADN của chúng. Có thể dùng cách tương tự đối với các quần thể. Gần đây, người ta đặc biệt chú ý đến sự phát tán của con người xuất phát từ châu Phi trong vài trăm nghìn năm trở lại đây. Sự phát tán của các bộ lạc qua Châu Âu, Châu Á, Châu Đại dương và Châu Mỹ trong và sau kỷ băng hà cuối cùng được xem xét một cách chi tiết với nhiều triển vọng đáng khích lệ.

IV.B TIẾN HÓA NHỜ CHỌN LỌC TỰ NHIÊN

Tiến hóa nhờ chọn lọc tự nhiên

Chọn lọc tự nhiên là quá trình mà môi trường hạn chế kích thước quần thể và bắt các cá thể phải cạnh tranh vì nguồn dinh dưỡng. Một số cá thể được di truyền những đặc điểm giúp chúng sống sót và sinh sản tốt hơn những cá thể khác. Kết quả là, bản sao của các gen quy định các tính trạng có lợi đó ngày càng trở nên phổ biến trong quần thể ở những thế hệ kế tiếp. Darwin đưa ra ví dụ tương tự quá trình chọn lọc trong tự nhiên là quá trình chọn lọc cây trồng và vật nuôi mà các nhà chọn giống tiến hành; họ giữ lại những cá thể có những đặc điểm tốt để nhân giống, loại bỏ những cá thể có đặc điểm kém hơn. Ông lấy chim bồ câu làm ví dụ. Thời đó, ông chưa hề biết quá trình di truyền diễn ra như thế nào; cũng như những người khác vào thời ông, ông hiểu không đúng về cơ chế kết hợp các đặc điểm từ bố, mẹ. Đây là hạn chế lớn nhất của Darwin, dù rằng ông phát hiện được nhiều dạng biến dị mới.

Chọn lọc tự nhiên tác động đến các biến dị đã tồn tại trong quần thể; ở đâu không có biến dị sẽ không có sự biến đổi. Tuy nhiên, các loài sẽ không bị tuyệt chủng chỉ vì lý do chúng kém thích nghi. Các loài chỉ bị tuyệt chủng khi môi trường sống bị mất hoặc bị biến đổi đến tình trạng mà không còn đủ số cá thể có kiểu gen thích nghi để có thể duy trì loài. Có rất nhiều loài động vật lớn ngày nay bị tuyệt chủng vì sự săn bắt của con người. Chúng thích nghi rất tốt và mang nhiều đặc điểm thích nghi tốt với môi trường. Khi động vật tuyệt chủng thì những đặc điểm đó và nguồn gen tương ứng của chúng cũng mất theo. Cả hai cùng mất đi vai trò sinh thái của mình, và chẳng còn gì để thích nghi nữa.

Những quan sát và kết luận của Darwin

Darwin đã có những quan sát và kết luận, có thể tóm tắt như sau:

Những quan sát:

Quan sát 1. Số cá thể của bất kỳ một quần thể nào cũng có thể tăng không giới hạn khi tất cả các cá thể cùng có điều kiện sống sót và sinh sản phù hợp. Bất kỳ loài nào cũng có thể nhanh chóng bao phủ toàn bộ trái đất nếu gặp các điều kiện sống thuận lợi.

Quan sát 2. Tiềm năng sinh sản, xét cho cùng, luôn bị hạn chế bởi nguồn sống. Ở tất cả các giai đoạn trong vòng đời của sinh vật đều có những nhân tố khác nhau hạn chế khả năng sống sót và sinh sản của chúng. Tuy vậy, tầm quan trọng của các nhân tố đó thay đổi theo thời gian, bao gồm cả trong vòng đời của từng cá thể sinh vật cũng như quá trình lịch sử giữa các thế hệ. Những nhân tố đó bao gồm: nguồn dinh dưỡng (thức ăn), hang ổ, địa dư, kích thước ổ, sự giao phối, ánh sáng mặt trời, vật săn mồi, bệnh tật và vật ký sinh.

Quan sát 3. Các cá thể không hoàn toàn giống nhau. Ở bất kỳ loài nào cũng có sự khác biệt (biến dị) giữa các cá thể. Những khác biệt đó ảnh hưởng một phần đến khả năng cạnh tranh, sống sót và sinh sản của chúng. Một số biến dị có thể di truyền được.

Các kết luận:

Kết luận 1. Từ quan sát 1 và 2, chắc chắn phải có sự cạnh tranh để sống sót và sinh sản. Trong quá trình đó, một số cá thể kém thành công hơn so với những cá thể khác.

Kết luận 2. Từ quan sát 3, một số đặc điểm khác biệt giữa các cá thể sẽ ảnh hưởng đến mức độ thành công của chúng trong cuộc đấu tranh để sống sót và sinh sản. Vì vậy, sự đào thải sẽ mang tính chọn lọc. Những cá thể mang những đặc điểm có lợi sẽ thành công hơn và sinh nhiều con cháu hơn những cá thể khác.

Kết luận 3. Những đặc điểm sinh học được quy định về mặt di truyền, nhờ vậy giúp các cá thể sống sót và sinh sản tốt trở nên phổ biến hơn ở thế hệ sau, làm biến đổi dần dần cấu trúc di truyền của quần thể qua thời gian.

Tiến hoá chỉ tác động đến phần biến dị được kiểm soát bởi các yếu tố di truyền. Những khác biệt do các yếu tố môi trường, như sự khác biệt về kích thước do được sinh ra ở nơi có nhiều thức ăn hoặc được chăm sóc tốt, chỉ có tác động đến sức đẻ con chứ không được truyền lại cho các thế hệ kế tiếp, vì thế sẽ không đóng góp đáng kể vào quá trình tiến hoá.

Darwin cho rằng thuật ngữ của Herbert Spencer "sự sống sót của những cá thể thích nghi nhất" không phải là cách giải thích tốt nhất về quá trình tiến hóa, vì giả thuyết này dễ gây nhầm lẫn về sự thích nghi vật lý. Khía cạnh quan trọng của sự thích nghi là việc sinh ra thế hệ con tốt hơn so với các cá thể khác trong loài. Các cá thể có thể thích nghi với sự biến đổi của môi trường (ví dụ như có bộ lông rậm hơn ở các vùng lạnh), nhưng đó không phải là sự tiến hoá vì nó không làm thay đổi cấu trúc di truyền của quần thể và không truyền được lại cho thế hệ sau. Tuy nhiên khả năng thích nghi thì lại được truyền lại cho thế hệ sau.

Phương thức chọn lọc

Có ba phương thức chọn lọc (*Hình 4.1*).

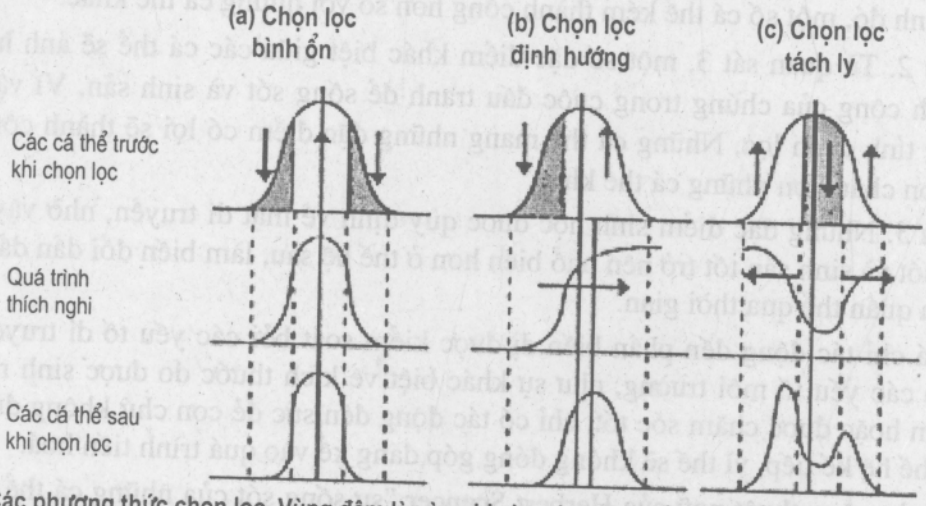
Chọn lọc bình ổn:

Chọn lọc bình ổn (*Hình 4.1a*) là quá trình loại bỏ những cá thể quá sai lệch so với dạng chuẩn và duy trì quần thể phân bố gần dạng tối ưu. Trong một môi

trường sinh thái ổn định thì hầu hết các cá thể đều thích nghi tốt, còn những dạng mới xuất hiện hoặc những dạng lệch lạc thường là kém thích nghi. Ví dụ, chim sẻ nhà (*Passer domesticus*) bị chết trong những trận bão tuyết dữ dội ở New York thường có cánh dài hơn hoặc ngắn hơn so với giá trị trung bình. Trẻ em có khối lượng sơ sinh cách xa 3,6kg thường bị chết hoặc bị tật bệnh ngay khi sinh hoặc sau khi sinh vì chúng hoặc quá to gây khó sinh đẻ, hoặc quá bé để sống sót. Ngày nay kiểu chọn lọc này giảm nhiều ở các nước có sự chăm sóc y tế trẻ sơ sinh tốt.

Chọn lọc định hướng:

Chọn lọc định hướng (*Hình 4.1b*) ưu tiên cho những cá thể có những đặc điểm ở một cực của phổ phân bố kiểu hình của quần thể. Ví dụ như loài chim sẻ Darwin, *Geospiza fortis*. Vào mùa khô ở quần đảo Galapagos, cỏ và những cây kích thước nhỏ thường sinh trưởng kém, tuy vậy những cây có rễ sâu và hạt to vẫn kết quả. Do vậy, chọn lọc ưu tiên những con chim sẻ có mỏ to, có khả năng ăn được những hạt lớn và chống lại những con chim có mỏ nhỏ, chỉ ăn được những hạt cỏ. Vào mùa mưa, khi có nhiều loại hạt nhỏ, những con chim này lại được bảo tồn. Chọn lọc định hướng tạo nên sự thay thế đặc điểm khi có hai loài cạnh tranh nhau. Nếu có hai loài sẻ giống nhau trên cùng một hòn đảo (cùng vùng phân bố) thì sẽ có xu hướng một loài được biệt hóa theo hướng ăn các hạt lớn và tiến hoá để có mỏ lớn, còn loài kia thì ngược lại. Hai loài sống riêng biệt (khác vùng phân bố) sẽ tiến hóa để có kích thước mỏ trung bình. Hiệu ứng này rất phổ biến ở những loài có quan hệ họ hàng.



Hình 4.1. **Các phương thức chọn lọc. Vùng đậm là vùng bị chọn lọc đào thải. Những mũi tên chỉ hướng biến đổi tần số kiểu hình theo thời gian. Đường liền, thẳng đứng là giá trị trung bình, các đường gạch nối để so sánh.**

Chọn lọc tách ly:

Chọn lọc tách ly (*Hình 4.1c*) chọn hai kiểu hình rất khác biệt nhau trong một quần thể. Ví dụ là sự chọn lọc dựa trên kích thước cơ thể (như ở chim cắt mái *Accipiter nisus* có kích thước lớn gấp đôi chim trống, chuyên bắt những con mồi lớn), hoặc chọn lọc dựa trên tập tính (như một số cá hồi đực trưởng thành sớm ở sông mà không cần phải bơi ra biển khi kích thước còn rất nhỏ; còn những con khác phải bơi ra biển, lớn lên rồi quay lại sông để sinh sản, chúng có kích thước lớn hơn, lấn át và thay thế những con đực có kích thước nhỏ và có khả năng thụ tinh nhiều trứng hơn, tuy vậy nhược điểm của chúng là vòng đời dài và phải ăn những con mồi lớn). Chọn lọc tách ly xảy ra ở quy mô một quần thể nếu có biến dị không liên tục với hai

dạng kiểu hình khác biệt rõ rệt (như giới tính và kích thước ở chim cắt). Tuy nhiên, có những dẫn chứng để tin rằng, đối với những tính trạng biến dị liên tục thì hai kiểu hình không thể cùng tồn tại nếu sự giao phối giữa chúng sinh ra những con lai có những kiểu hình trung gian không thích nghi (xem mục III.G). Sau đó, kiểu hình ít phổ biến hơn trong số hai kiểu hình có thể bị đào thải (xem mục IV.H). Chọn lọc tách ly có thể làm cho một loài tiến hoá thành hai loài mới, mỗi loài mới thích nghi với một giá trị tối ưu. Tuy nhiên, hiện tượng đó cần có cơ chế cách ly trước giao phối để ngăn cản sự giao phối tiếp tục giữa hai dạng.

Hình thái loài khởi sinh

Darwin coi các chủng là các loài khởi sinh. Có sự phân bố liên tục giữa các chủng chỉ khác nhau chút ít với các loài đã phân ly hoàn toàn. Sự phân ly xảy ra khi hai quần thể trải qua sự biến đổi di truyền trong một thời gian dài. Hiện tượng đó có thể do chọn lọc định hướng theo hai kiểu hình khác nhau ở hai quần thể. Thậm chí, nếu chọn lọc cùng theo một hướng ở cả hai quần thể (ví dụ, kích thước mỏ ở chim, nhiệt độ tối ưu...) thì ở mỗi quần thể, vẫn có thể là các tổ hợp alen khác nhau được chọn lọc. Ví dụ, giả sử hai gen *A* và *B* mã hoá cùng một loại enzym và chọn lọc xảy ra theo hướng tạo ít enzym hơn. Khi đó kiểu gen *AABB* sẽ sinh ra quá nhiều enzym và như vậy gen *A* hoặc *B* sẽ bị mất đi. Kiểu gen *AAbb* có thể sẽ được cố định ở một quần thể, còn ở quần thể kia là *aaBB* (*aabb* sẽ có hiệu ứng gây chết vì không sinh ra enzym). Còn có những sai khác xuất hiện do ngẫu nhiên (đột biến hoặc lạc dòng di truyền). Khi hai quần thể đã khác biệt rõ nét về hình dạng, tập tính, sinh thái hoặc di truyền, chúng có thể được coi là hai loài riêng biệt.

Chọn lọc giới tính

Quá trình chọn lọc có thể có ảnh hưởng khác nhau đến hai giới tính. Sở dĩ như vậy bởi vì ở nhiều loài, các tiểu giao tử đực (tinh trùng và hạt phấn) thường có kích thước nhỏ, được tạo ra với số lượng lớn còn các đại giao tử (trứng và noãn) thường chỉ được tạo ra với kích thước lớn và số lượng ít; các cá thể cái ngoài chức năng sinh sản còn phải nuôi con. Do vậy, các cá thể đực có cơ hội để tạo ra nhiều con cái hơn. Sự **chọn lọc nội giới tính** là mối quan hệ cạnh tranh mãnh liệt giữa những cá thể của cùng một giới tính (thông thường là giữa các con đực để tranh giành con cái), và kết quả thường là chọn ra những cá thể có kích thước lớn hoặc vũ khí bảo vệ (như sừng) hiệu quả hơn khi chúng giao chiến với nhau. Sự khác biệt thấy rõ nhất ở loài Hải cẩu voi (elephant seal) với những con đực có phần thân sau lớn. Trong một số ít trường hợp, có sự cạnh tranh giữa những con cái để "giành giật" con đực, như ở các loài chim làm tổ ngoài không gian mở. Các con cái của một số loài chim nước thuộc chi *Jacana* bảo vệ lãnh thổ của nó với một số con chim đực, trong đó mỗi con đực bảo vệ một ổ trứng của con cái. Sự **chọn lọc liên giới tính** là sự cạnh tranh giữa các cá thể của cùng một giới tính (thường là giới tính đực) để hấp dẫn "bạn tình" và liên quan đến mối quan hệ giữa các cá thể của giới tính kia. Điều này thường dẫn đến sự chọn lọc các tính trạng hình thức như đuôi công hay bờm, được xem như dấu hiệu để tìm những cá thể đực phù hợp với cá thể cái. Các các thể cái lúc đó được gọi là **giới kén chọn**, còn sự trang trí hình thức ở con đực được coi như một **sự phô trương**. Ở một số nhóm, sự phô trương này không rõ, chủ yếu là nhằm mục đích hấp dẫn con cái, hay để đe dọa những con đực khác, hay vì cả hai mục đích.

ADN/gen ích kỷ

Hệ gen của các sinh vật sinh sản hữu tính bị tái tạo trong quá trình phát sinh giao tử do sự tái tổ hợp và sự sắp xếp lại. Hệ gen của các cá thể chỉ là tạm thời; chỉ các bản sao các gen và ADN được truyền qua các thế hệ. Hiện tượng đó dẫn đến quan điểm về "ADN và gen ích kỷ", bởi vì những đoạn ADN hoặc những gen nào tăng cường được khả năng truyền các bản sao của chúng cho thế hệ sau mới có ưu thế chọn lọc so với các gen khác. Gen ích kỷ có thể là gen quy định tập tính sẵn sàng hy sinh ở cá thể này để xúc tiến quá trình sinh sản mà cá thể khác thực hiện để truyền các bản sao của các gen đó cho thế hệ sau. Điều đó có thể minh họa trên ví dụ ở ong và kiến – những con vật đã giúp đỡ cho chị, em chúng trở thành những con chúa để thực hiện chuyên hóa chức năng sinh sản. Những con đực của nhóm động vật này (*Hymenoptera*) là đơn bội, còn những con cái lưỡng bội được di truyền đủ các gen từ bố của chúng. Như vậy, khoảng 75% hệ gen của chúng liên quan với nhau chứ không phải 50% như thường lệ. Như vậy, các gen sẽ được truyền cho đời con hiệu quả hơn thông qua con cái. Nhiều loài côn trùng xã hội tiến hoá theo cách này.

Các yếu tố di truyền vận động là các đoạn ADN ngắn có thể được sao chép và xen vào các vị trí khác nhau trên các nhiễm sắc thể. Có lẽ chúng là các ADN "ích kỷ". Sự chuyển vị của chúng có thể được thực hiện nhờ các enzym do chính các yếu tố di truyền vận động mã hóa (giống như các retrovirus) hoặc do các protein được mã hóa bởi các trình tự ADN nằm đâu đó trong hệ gen. Nếu một đoạn ADN nào đó có khả năng tăng cường sự vận động của nó vào các giao tử và hợp tử với tần số cao hơn tần số mong đợi theo quy luật Mendel thì chúng có ưu thế chọn lọc lớn. Hiện tượng này được gọi là sự điều khiển giảm phân. Trong một số trường hợp nó có thể làm giảm khả năng sinh sản của cá thể. Ví dụ: alen *SD* (*s*egregation *d*istorter) ở *Drosophila* có thể phá huỷ các tinh trùng không mang các bản sao của chính nó. Khi đó nó là một ADN ký sinh ích kỷ thực sự. Còn có các nhiễm sắc thể ký sinh kiên trì tồn tại trong các tế bào mầm nhưng bị mất đi ở các tế bào soma, nơi mà chúng có thể làm giảm giá trị chọn lọc của cá thể. Có những nhiễm sắc thể ký sinh khác xâm nhập được vào nhân tế bào trứng và tránh được các tế bào thể cực.

Điểm lại lịch sử

Quan điểm tiến hoá nhờ sự biến đổi dần dần qua các thế hệ kế tiếp đã được chấp nhận khi **Darwin** đưa ra luận điểm **chọn lọc tự nhiên là cơ chế tiến hoá** trong cuốn "Nguồn gốc các loài" của mình xuất bản vào năm 1859. Darwin hình thành quan điểm này vào tháng 9 năm 1838, song nó chỉ được công bố và đăng tải vào năm 1858, sau khi **Alfred Russel Wallase**, người đã một cách độc lập có những kết luận tương tự. Các thuyết tiến hoá trước đó đều dựa trên quan niệm cổ xưa về trật tự sinh giới, gồm chất sống dạng nhày ở đáy, còn con người, các thiên thần và chúa trời ở trên đỉnh; mọi sinh vật đều cố gắng để trở nên giống với các thiên thần. **Larmark** đã cố gắng trình bày quan điểm đó như một giả thuyết khoa học, trong đó cho rằng những tiến bộ mà bố, mẹ cố gắng có được sẽ được truyền lại cho con cái chúng (**sự di truyền các tính trạng tập nhiễm**). Bằng chứng cho thấy quan điểm đó không chính xác, chẳng hạn như việc cắt cụt đuôi liên tục của các con chuột khi mới sinh qua nhiều thế hệ cũng không thể tạo ra được một dòng chuột không đuôi.

IV.C TRẠNG THÁI CÂN BẰNG HARDY-WEINBERG

Mở đầu

Di truyền học quần thể là môn nghiên cứu các alen của các gen trong quần thể và các dạng áp lực duy trì hoặc biến đổi tần số của các alen và các kiểu gen cụ thể trong quần thể. Một quần thể liên phối địa phương được gọi là deme để phân biệt nó với quần thể địa lý. Có thể rất khó xác định chính xác deme, ví dụ những người sống trong cùng một vùng có thể chọn bạn đời dựa trên các tiêu chuẩn về tôn giáo, chủng tộc, sức khoẻ và tầng lớp xã hội chứ không phải hoàn toàn ngẫu nhiên. Tương tự, những côn trùng sống trên cùng một cây trong rừng có thể chỉ giao phối với nhau mà không bay sang những cây khác để giao phối. Vậy quần thể có thể là loài côn trùng của toàn bộ khu rừng đó, hoặc là mỗi cây lại có một quần thể bán cách ly của riêng mình. Các quần thể tự nhiên gồm rất nhiều cá thể liên phối với nhau, mỗi cá thể có tổ hợp gen riêng, nhưng quần thể cùng chia sẻ một vốn gen chung. Tiến hoá là sự biến đổi tần số các alen của vốn gen, vì thế mà việc nghiên cứu di truyền học quần thể có ý nghĩa quan trọng. Một số kiểu gen, ví như các kiểu gen liên quan đến các bệnh di truyền ở người, thường hiếm và di truyền học quần thể giúp ta tìm hiểu vì sao và tiên đoán tỷ lệ mắc các bệnh đó. Hầu hết các kiểu hình đều gần với giá trị trung bình, nhưng có một số cá thể mang các tổ hợp gen hiếm làm cho kiểu hình của chúng sai khác rõ ràng với các cá thể bình thường.

Hardy và Weinberg, độc lập với nhau, đã cùng công bố những tính toán cho thấy những tính trạng không mong muốn không thể trở nên phổ biến hơn trong quần thể vì chúng nằm trong các tổ hợp hiếm gặp so với các alen phổ biến hơn. Họ chứng minh rằng, tần số của những kiểu gen cụ thể trong quần thể của các loài sinh vật lưỡng bội sinh sản hữu tính đạt trạng thái cân bằng ngay sau một thế hệ ngẫu phối và không thay đổi cho đến khi có nhân tố nào đó làm thay đổi tần số alen của quần thể. Trạng thái cân bằng Hardy-Weinberg cần cơ chế ngẫu phối, không có chọn lọc hoặc di nhập... tác động đến các alen và các kiểu gen.

Mức độ phù hợp của tần số các kiểu gen quan sát được với tần số lý thuyết tính theo cân bằng Hardy-Weinberg có thể kiểm nghiệm được bằng thống kê theo phương pháp χ^2. Mức độ sai lệch giữa hai số liệu trên là bằng chứng cho thấy có sự tác động của các yếu tố như chọn lọc, di nhập, giao phối chọn lựa, lạc dòng di truyền...

Tính tần số alen

Tần số alen (giả sử *A*) trong quần thể là số lượng alen *A* chia cho tổng số alen của locut gen đó. Ở một loài lưỡng bội, một quần thể có *N* cá thể sẽ sinh ra *2N* giao tử và có *2N* alen ở mỗi locut gen. Giả sử một locut gen có 2 alen, *A* và *a*, thì số lượng alen *A* sẽ bằng 2 lần số lượng đồng hợp tử *AA* cộng với số lượng dị hợp tử *Aa* (vì mỗi đồng hợp tử có 2 alen *A* và mỗi dị hợp tử có 1 alen *A*). Tần số alen *A* bằng số lượng alen *A* chia cho tổng số alen là *2N*. Tần số alen *a* cũng được tính tương tự như vậy. Nếu ký hiệu *n* là số lượng, *A* và *a* là các alen, *AA*, *Aa* và *aa* là các kiểu gen, thì ta có:

$$n_A = 2n_{AA} + n_{Aa}$$

Nếu gọi *p* là tần số alen *A*, thì

$$p = \frac{N_A}{2N} = \frac{2N_{AA} + N_{Aa}}{2N}$$

và q là tần số alen a, thì

$$q = \frac{N_a}{2N} = \frac{2N_{aa} + N_{Aa}}{2N}$$

Chú ý rằng $n_a + n_A = 2N$, và $p + q = 1$.

Với một gen có hơn 2 alen thì cách tính cũng tương tự như vậy, và tần số alen thứ ba thường được ký hiệu là r.

Tần số alen chỉ có thể xác định chính xác ở quần thể lớn và khi phân biệt được các thể đồng hợp tử và dị hợp tử. Thường khó phân biệt các thể dị hợp tử trong trường hợp trội hoàn toàn, tuy nhiên, ngày nay kỹ thuật phân tích chỉ thị ADN, kể cả kỹ thuật PCR đã giúp giải quyết vấn đề này.

Ví dụ: Hai alen quy định nhóm máu MN ở người nằm trên nhiễm sắc thể số 2 là L^M và L^N (thường gọi đơn giản là alen *M* và *N*) (*Bảng 4.1*). Hai alen này đồng trội, vì thế có thể xác định thể dị hợp tử bằng cách thử kháng thể. Trong một quần thể gồm 1000 người Mỹ da trắng có 357 người *MM*, 485 người *MN* và 158 người *NN*. 357 người *MM* có 714 alen *M*, 485 người *MN* có 485 alen *M*, nên tần số alen *M* (=*p*) là

$$p = \frac{(2 \times 357) + 485}{2000} = 0,5995$$

Bảng 4.1. **Cách tính tần số alen M và N trong quần thể**

	Kiểu gen (= kiểu hình)			
	MM	*MN*	*NN*	Tổng số
Số cá thể	357	485	158	1000
Số alen *M*	714	485	0	1199
Số alen *N*	0	485	316	801
Tổng số	715	970	316	2000

Tần số alen *M*: 1199/2000 = 0,5995; tần số alen *N* = 801/2000 = 0,4005.

Tần số các kiểu gen ở trạng thái cân bằng

Trạng thái cân bằng sẽ đạt được sau một thế hệ ngẫu phối. Tần số kết hợp giữa giao tử đực mang alen *A* (có tần số *p*) giao phối với giao tử cái mang alen *A* (cũng có tần số *p*) sẽ là $p \times p = p^2$. Tương tự, tần số để giao tử đực mang alen a (với tần số *q*) kết hợp với giao tử cái mang alen a (cũng với tần số *q*) là $q \times q = q^2$.

Dị hợp tử *Aa* được tạo ra do sự kết hợp của các giao tử mang các alen khác nhau. Tần số để giao tử đực mang alen *A* kết hợp với giao tử cái mang alen *a* là $p \times q = pq$, và tần số để giao tử đực mang alen *a* kết hợp với giao tử cái mang alen *A* là $q \times p = pq$, nên tổng tần số dị hợp tử *Aa* là $2pq$.

Các tần số trên đây là kết quả triển khai nhị thức:

$$(p + q) \times (p + q) = (p + q)^2 = p^2 + 2pq + q^2$$

Ví dụ về tần số các nhóm máu *MM*, *MN* và *NN* ở những người Mỹ da trắng (*Hình 4.2a*) và những người thổ dân Úc (*Hình 4.2b*) minh hoạ cho cách tính nói trên. Chú ý rằng tần số các alen rất khác nhau ở hai quần thể này.

Tần số đồng hợp tử bằng bình phương tần số các alen tương ứng, còn tần số dị hợp tử bằng hai lần tích tần số của hai alen. Cách tính này cũng áp dụng cho một gen có nhiều hơn hai alen. *Hình 4.3* là cách tính tần số 6 kiểu gen của nhóm máu ABO ở người.

a)

Vốn gen trứng \ Vốn gen tinh trùng	Tần số alen $M = p = 0,6$	Tần số của $N = q = 0,4$
Tần số alen $M = p = 0,6$	*MM* $p^2 = 0,36$	*MN* $pq = 0,24$
Tần số $N = q = 0,4$	*MN* $pq = 0,24$	*NN* $q^2 = 0,16$

b)

Vốn gen trứng \ Vốn gen tinh trùng	Tần số alen $M = p = 0,178$	Tần số của $N=q=0,822$
Tần số alen M $p=0,178$	*MM* $p^2=0,0317$	*MN* $pq = 0,146$
Tần số $N = q = 0,822$	*MN* $pq=0,146$	*NN* $q^2 = 0,676$

Hình 4.2. **Tần số các nhóm máu *MM*, *MN* và *NN* ở người Mỹ da trắng (a) và thổ dân Châu Úc (b). Alen L^M với tần số p mã hoá cho kháng nguyên *M*, alen L^N với tần số q mã hoá cho kháng nguyên *N*. Hai alen đồng trội, vì vậy, kiểu gen L^ML^M quy định nhóm máu *M*, L^NL^N quy định nhóm máu *N* và kiểu gen L^ML^N quy định nhóm máu *MN*.**

Những điểm cần lưu ý:

(i) Khi tần số một alen (giả sử q) thấp thì tần số đồng hợp tử q^2 rất thấp. Ví dụ như tần số đồng hợp tử *MM* ở thổ dân Úc (*Hình 4.2b*) và tần số kiểu gen I^BI^B (*Hình 4.3*).

(ii) Alen hiếm thường tồn tại ở trạng thái dị hợp tử (ví dụ alen *M* ở thổ dân Úc). Nếu đó là alen lặn, nó sẽ không được biểu hiện và không thể xác định được, vì thế cả chọn lọc tự nhiên cũng như các chương trình ưu sinh học không thể tác động đào thải nó.

Tần số của các alen

	$I^A = p = 0,548$	$i = r = 0,3995$	
I^A p	I^AI^A A $p^2 = 0,3$	I^Ai A $pr = 0,219$	I^B $q = 0,0525$ / I^{AB} AB $pq = 0,029$
i r	I^Ai A $p^2 = 0,219$	ii O $r^2 = 0,16$	I^B B $qr = 0,021$
	AB $pq = 0,029$	B $qr = 0,021$	I^BI^B B $q^2 = 0,0028$

Nhóm máu	Kiểu gen	Tần số	
A	I^AI^A I^Ai	$p^2 = 0,3$ $2pr = 0,438$	**0,738**
B	I^BI^B I^Bi	$q^2 = 0,0028$ $2qr = 0,042$	**0,448**
AB	I^AI^B	$2pq = 0,058$	**0,058**
O	ii	$r^2 = 0,16$	**0,16**

Hình 4.3. **Tần số các nhóm máu *A*, *B*, *AB* và *O* theo cân bằng Hardy-Weinberg. Tần số alen I^A là p, I^B là q và i là r. I^A, I^B mã hoá các enzym sinh ra các kháng thể của nhóm máu *A* và *B* và đồng trội, còn i là alen đột biến lặn, không có hiệu ứng.**

Tác nhân gây sai lệch trạng thái cân bằng

Nếu một quần thể không ở trạng thái cân bằng Hardy-Weinberg thì có thể do một số nguyên nhân sau:

– **Chọn lọc** tác động đến khả năng sống sót và độ hữu thụ của những kiểu gen nhất định ở thế hệ bố, mẹ sẽ làm sai lệch tần số các kiểu gen ở hợp tử. Sức sống khác biệt của các kiểu gen sẽ làm sai lệch tần số các kiểu gen ở các cá thể trưởng thành. Trong trường hợp này, sự sai khác so với trạng thái cân bằng càng lớn khi mẫu chọn bao gồm các cá thể có tuổi trưởng thành càng cao.

– **Sự nhập cư** sẽ mang đến các kiểu gen xác định từ một vốn gen khác, gây ra sai lệch. Quần thể hỗn hợp sẽ luôn sinh ra nhiều thể đồng hợp tử. Nếu các kiểu gen khác nhau có xu hướng di cư hoặc nhập khác nhau thì mức độ sai lệch sẽ theo một chiều hướng nhất định.

– **Giao phối chọn lựa.** Nếu các cá thể chọn cặp giao phối với kiểu gen tương tự thì sẽ sinh ra quá nhiều đồng hợp tử (ví dụ *AA* x *AA* sẽ sinh ra toàn *AA*), ngược lại nếu chọn lựa theo kiểu gen khác biệt, sẽ chỉ sinh ra các thể dị hợp tử (*AA* x *aa* sinh ra *Aa*).

– **Sự phân nhỏ quần thể** sẽ làm tăng tần số các thể đồng hợp tử. Hiện tượng này xuất hiện nếu mức độ vận động của sinh vật và hiện tượng liên phối nhỏ hơn nhiều so với khu vực lấy mẫu, vì khi đó các cá thể chỉ có vốn gen để giao phối rất nhỏ và chúng không thể giao phối ngẫu nhiên. Do hiệu ứng đó mà các kiểu gen nền của các quần thể khác nhau luôn có xu hướng tăng tần số đồng hợp tử. Cấu trúc xã hội của loài người cũng có hiệu ứng tương tự. Trạng thái cực đoan là sự nội phối giữa các thành viên của cùng một dòng họ. Hiện tượng đó được minh chứng bằng tần số các dị tật di truyền hiếm gặp ở người. Ví dụ, 33% ca bệnh *alkapton niệu* và 54% ca bệnh *đầu nhỏ* là do có cặp bố, mẹ là anh em họ hàng. Điều đó cho thấy, ở người, tần số đồng hợp tử hiếm sẽ tăng đáng kể do sự giao phối có chọn lựa của các cá thể có quan hệ họ hàng. Tần số đồng hợp tử biến đổi từ giá trị q^2 theo ngẫu phối đến *pq/16* cho trường hợp kết hôn giữa các anh em họ.

– **Đột biến** phá huỷ các alen đang tồn tại và sinh ra các alen mới. Do tần số đột biến thấp, quá trình này tác động chậm.

– **Hiện tượng lạc dòng di truyền** hay cơ hội di truyền thành công khác nhau của các alen nhất định từ thế hệ này sang thế hệ khác sẽ gây ra sự sai khác nhỏ về tần số kiểu gen so với tần số mong đợi. Thuật ngữ "lạc dòng di truyền" chỉ áp dụng cho trường hợp thay đổi ngẫu nhiên tần số alen giữa các thế hệ chứ không phải cho tất cả các trường hợp sai lệch so với cân bằng Hardy-Weinberg.

Đánh giá độ phù hợp

Phép thử χ^2 được dùng để so sánh hai mẫu dựa trên hai hoặc nhiều chỉ tiêu. Giá trị χ^2 càng nhỏ thì mức độ phù hợp càng cao. Để kiểm tra độ phù hợp với cân bằng Hardy-Weinberg, ta dùng công thức:

$$\chi^2 = \sum \frac{(\text{tần số quan sát} - \text{tần số lý thuyết})^2}{\text{tần số lý thuyết}}$$

Nếu có hai alen, sẽ có một bậc tự do (bằng số alen trừ đi một). Cách tính độ phù hợp trong ví dụ về nhóm máu *MN* ở những người Mỹ da trắng trên đây được trình bày ở *Bảng 4.2*. Giả thuyết là tần số của mẫu không sai khác với tần số lý thuyết theo Hardy-Weinberg. Giá trị

χ^2 thu được là 0,4236. Giá trị này rất nhỏ, có nghĩa là chỉ có sai khác nhỏ giữa số liệu quan sát và số liệu lý thuyết.

Bảng 4.2. **Kiểm tra độ phù hợp của các số liệu quan sát về nhóm máu MN so với trạng thái cân bằng Hardy-Weinberg bằng phương pháp $\chi 2$**

	Kiểu gen			
	MM	MN	NN	Tổng số
Số liệu quan sát	357	485	158	1000
Số liệu lý thuyết	$1000p^2$	1000 x $2pq$	$1000q^2$	
	359,4	480,2	160,4	1000
SL quan sát – SL lý thuyết	-2,4	4,8	-2,4	
(SL quan sát – SL lý thuyết)^2	5,76	23,3	5,76	
(SL quan sát – SL lý thuyết)^2/SL lý thuyết	0,016	0,0485	0,3591	

SL: số liệu. Tần số alen $M = p = 0,5995$, tần số alen $N = q = 0,4005$; $\chi^2_{(1\ bậc\ tự\ do)} = 0,16 + 0,0485 + 0,03591 = 0,4236$.

Sử dụng bảng giá trị χ^2, chúng ta có thể thấy rằng, rất có thể chúng ta còn thu được độ lệch do ngẫu nhiên còn lớn hơn độ lệch tính được trên đây, vì thế chúng ta chấp nhận giả thuyết trùng khớp, nghĩa là tần số thu được phù hợp với tần số lý thuyết. Giá trị χ^2 (với bậc tự do là 1) lớn hơn 7,879 cho thấy xác suất phù hợp do ngẫu nhiên chỉ có 5%, 95% còn lại là sai lệch thực sự so với trạng thái cân bằng Hardy-Weinberg. Chú ý rằng, độ sai lệch lớn so với lý thuyết sẽ cho giá trị χ^2 lớn và ngược lại.

Ước tính tần số alen

Trong nhiều trường hợp (ví dụ đối với các alen lặn gây bệnh ở người), ta dễ dàng xác định các thể đồng hợp tử về alen lặn (những người bị bệnh), song khó xác định các thể dị hợp tử và đồng hợp tử trội (những người bình thường). Trong trường hợp đó, tần số alen lặn có thể được ước tính dựa vào trạng thái cân bằng Hardy-Weinberg. Giả sử quần thể ở trạng thái cân bằng thì tần số đồng hợp tử lặn phải bằng q^2, vậy:

$$q = \sqrt{q^2} = \sqrt{\frac{\text{Số cá thể đồng hợp tử lặn}}{\text{Tổng số cá thể}}}$$

Ví dụ, khoảng 1/2500 người da trắng bị bệnh xơ nang khi sơ sinh. Họ là đồng hợp tử về alen không hoạt động chức năng của gen quy định protein vận chuyển chloride tham gia vào quá trình tiết chất nhày. Không có chức năng này, chất nhày rất đặc và tích tụ lại, phá huỷ tuyến tuỵ, ruột non và phá hoại phổi một cách nghiêm trọng. Vì tần số đồng hợp tử là $q^2 = 1/2500 = 0,0004$, nên $q = 0,02$, hay1/50.

Trong thực tế, phương pháp ước tính này có thể có sai số vì nhiều lý do: (i) hầu hết các quần thể đều có mức độ nội phối nhất định. Con người có thể lấy người có quan hệ họ hàng, hoặc chọn bạn đời theo quan hệ láng giềng, chủng tộc, tôn giáo, trình độ học vấn, thậm chí

chiều cao... làm tăng tần số đồng hợp tử; (ii) tần số đồng hợp tử (những người bị bệnh) rất thấp, nên dễ có sai sót đáng kể khi tính toán. Ví dụ, tần số thực sự là 4 người bị bệnh trong một quần thể gồm 100.000 người, vậy một nghiên cứu ngẫu nhiên có thể tìm được con số lớn hơn hoặc nhỏ hơn thực tế là 2 người (2 người hoặc 6 người bị bệnh), làm cho tần số ước tính tăng hoặc giảm 2 lần; (iii) một số bệnh di truyền có mức ngoại hiện thấp, nghĩa là không phải mọi cá thể đồng hợp tử đều biểu hiện bệnh. Điều đó là do sự biểu hiện của bệnh đôi khi còn phụ thuộc một số điều kiện khác nữa.

Chúng ta có thể tính tần số thể dị hợp tử (thể mang) về bệnh xơ nang:

$$2pq = 2 \times (1 - 0{,}02) \times 0{,}02 = 0{,}0392 \text{ (xấp xỉ 1/25)}$$

Nói cách khác, cứ 25 người da trắng có một người mang alen lặn gây chết này. Tần số này là khá cao. Lời giải thích cho hiện tượng này có thể là: những thể mang gen gây bệnh xơ nang lại tăng khả năng chịu đựng với chất độc của vi trùng tả - tác chất gây chết do làm mất nước trong ruột (họ còn có khả năng chịu đựng cao hơn đối với một số loài *Salmonella*). Những con chuột chuyển gen dị hợp tử về gen xơ nang có tốc độ mất nước chỉ bằng một nửa so với các con đồng hợp tử khi xử lý bằng chất độc này, còn những con chuột đồng hợp tử về gen xơ nang thì hoàn toàn không bị mất nước. Đó là một ví dụ về trạng thái cân bằng do chọn lọc theo ưu thế dị hợp tử, giống như các thể mang gen gây bệnh hồng cầu hình liềm đối với khả năng nhiễm **bệnh sốt rét.**

IV.D ĐA DẠNG DI TRUYỀN

Mở đầu

Trong quần thể, hầu hết các gen đều tồn tại ở trạng thái có vài alen. Trên thực tế trước đây, hạn chế về khả năng phát hiện các biến dị là do nguyên nhân về kỹ thuật. Kỹ thuật giải trình tự ADN gần đây cho phép phát hiện những khác biệt trong trình tự ADN, kể cả các đột biến không gây nên sự thay đổi kiểu hình. Những đột biến đó có thể là do không làm thay đổi amino acid được mã hoá (ví dụ, bộ ba AAA và AAG cùng mã hoá lysin), vì thế không ảnh hưởng đến chuỗi polypeptid hoặc sản phẩm protein, hoặc không làm thay đổi hoạt động chức năng của phân tử protein. Những đột biến này được gọi là các **đột biến câm**. Tuy nhiên, các bộ ba thay thế thường không được sử dụng với mức độ tương ứng cho thấy: trên ribosom, có thể bộ ba này được dịch mã hiệu quả hơn bộ ba khác. Biến dị về trình tự còn tồn tại ở cả các vị trí không mã hoá, kể cả trường hợp có hoặc không có vai trò điều hòa. Chỉ những vùng trình tự quan trọng là chịu tác động của chọn lọc. Người ta cho rằng, trong hầu hết các quần thể vốn không có quá nhiều biến dị là do kết quả của chọn lọc; nếu đúng như vậy thì sẽ có nhiều cá thể phải chết để chọn lọc có thể tác động và duy trì biến dị. Cơ chế sản sinh và duy trì khả năng biến dị có thể được xem xét để xác định mức độ tác động của các yếu tố chọn lọc, đột biến và lạc dòng di truyền đến sự hình thành và duy trì các biến dị.

Nguồn biến dị

Mọi biến dị di truyền đều do đột biến - những biến đổi về trình tự base trên ADN (hoặc ARN của một số loại virus). Những biến đổi đó có thể làm hỏng khả năng biểu hiện chức năng bình thường của một

alen, đồng thời có hại cho giá trị thích nghi của kiểu gen. Đột biến vô hiệu gây ra do sự phá huỷ hoàn toàn chức năng của một gen, thường là do mất đoạn. Đột biến tạo ra các bộ ba kết thúc chuỗi (đột biến vô nghĩa) hoặc các đột biến dịch khung xảy ra ở các bộ ba đầu của trình tự mã hoá, hoặc các đột biến điểm làm biến đổi các amino acid quan trọng (như các amino acid ở vị trí hoạt hoá của các enzym) đều có hiệu ứng tương tự. Các alen mất chức năng đều có thể coi là một loại. Chúng là các alen lặn nếu chỉ một bản sao duy nhất của alen hoạt động chức năng là đủ để sinh ra kiểu hình bình thường. Ở người, chúng thường chỉ gây hại khi ở trạng thái đồng hợp tử và có thể gây ra các bệnh di truyền. Một số đột biến vô hiệu cũng có thể có lợi (như việc mất alen I^A ngăn cản sự biểu hiện của kháng nguyên nhóm máu A lại giúp cho cá thể trở nên có khả năng kháng lại bệnh đậu mùa ở mức độ nhất định).

Đột biến điểm làm biến đổi chỉ một amino acid trên chuỗi polypeptid có thể gây hiệu ứng kiểu hình không nghiêm trọng. Chúng có thể có hại một chút, song không đủ để gây hậu quả nghiêm trọng cho cá thể mang đột biến. Các đột biến đó được gọi là trung tính vì chọn lọc không tác động đến tần số của nó. Một alen có thể được gọi là trung tính nếu tần số của nó được kiểm soát chủ yếu bởi hiện tượng lạc dòng di truyền chứ không phải bởi chọn lọc. Nằm giữa hai loại đột biến đó là các đột biến có mức độ gây hại khác nhau. Rất hiếm có những đột biến làm tăng giá trị chọn lọc của kiểu gen. Những đột biến có ưu thế như vậy sẽ được chọn lọc và có thể thay thế các alen đang tồn tại.

Chọn lọc chống lại đột biến

Đột biến không ngừng tạo ra các alen mất chức năng. Chọn lọc tác động chống lại chúng, nhưng hầu hết chúng ở trạng thái lặn nên chọn lọc chỉ chống lại được các cá thể đồng hợp tử. Sự gia tăng tần số alen đột biến làm tăng tần số các đồng hợp tử so với dị hợp tử, và chọn lọc tăng cường chống lại các đồng hợp tử, ngăn cản sự gia tăng tiếp tục tần số alen đột biến. Ở trạng thái cân bằng, số đột biến mất chức năng mới sinh ra cân bằng với số bị mất đi do chọn lọc đào thải. Đơn giản hơn, chúng ta giả sử rằng, các alen mất chức năng là lặn, vì thế chọn lọc chỉ chống lại các cá thể đồng hợp tử có tần số q^2. Trong một quần thể lưỡng bội có N cá thể sẽ có $2N$ alen. Giả sử tần số alen đột biến q thấp, mỗi thế hệ sẽ có $2pN\mu$ đột biến mới (với μ là tốc độ đột biến trên một thế hệ và p là tần số alen hoạt động chức năng), và $2Nsq^2$ alen đột biến bị chọn lọc đào thải (với s là hệ số chọn lọc chống lại đồng hợp tử và s = 1 đối với alen lặn gây chết).

Ở trạng thái cân bằng, $2pN\mu = 2Nsq^2$

p thường được coi bằng 1, trong trường hợp đó, chia hai vế cho $2N$, ta có

$$\mu = sq^2, \text{ hay } q^2 = \mu/s \text{ và } q = \sqrt{\mu/s}$$

Nếu không có hiện tượng trội thì dị hợp tử cũng chịu tác động của chọn lọc với hệ số $0{,}5s$, khi đó $q = 2\mu/s$.

Lạc dòng di truyền

Kiểu gen của bất kỳ giao tử hoặc hợp tử nào cũng phụ thuộc vào sự kiện là nó được di truyền nhiễm sắc thể nào từ cặp tương đồng của bố, mẹ. Mỗi alen đều có xác suất được truyền cho con cụ thể là 1/2. Nếu cho rằng, trung bình hai bố, mẹ (lưỡng bội) sinh hai con thì mỗi alen sẽ được truyền trung bình một lần, mỗi bản sao của thế hệ bố, mẹ sẽ có mặt một lần ở thế hệ con, giống như ta tung đồng xu, xác suất để có mặt sấp và mặt ngửa là như nhau. Tuy nhiên, tần số các alen ở cá thể

có thể cao hoặc thấp, giống như ta tung nhiều đồng xu sẽ thu được số mặt sấp, ngửa khác nhau ở mỗi đồng xu. Như vậy, vốn gen sẽ bị biến đổi ở thế hệ sau, và quá trình cứ lặp lại như vậy làm cho tần số alen bị biến đổi theo thời gian. Phương sai tần số alen được tính theo phân phối nhị thức. Phương sai S^2 của tần số q của alen a được tính bằng $S^2_q = p_0q_0/2N$ (với p_0 và q_0 là tần số khởi đầu của alen A và a). Với t thế hệ

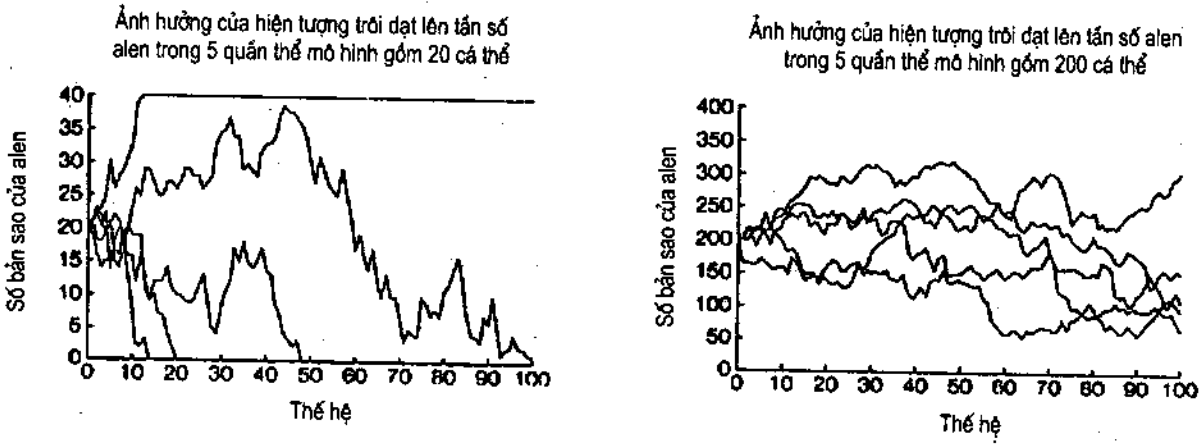
$$S^2_q = p_0q_0\{1 - [1 - 1/2N]^t\}$$

Có ba điểm cần lưu ý:

(i) Lạc dòng di truyền phụ thuộc vào tần số alen ban đầu. Các alen hiếm thường rất nhạy với hiện tượng lạc dòng di truyền. Chẳng hạn, nếu một alen hiếm chỉ có ở một con cá tuyết, nó có thể được truyền cho nửa triệu trứng, 1000 trứng trong số đó có thể sống sót, hoặc con cá đó có thể bị ăn thịt trước khi nó sinh sản và alen đó bị đào thải. Càng nhiều cá thể mang alen đó thì xác suất để alen đó được truyền lại cho thế hệ sau càng gần với 1 (100%).

(ii) Đây là quá trình cộng gộp qua thời gian, tuy nhiên tần số alen có thể tăng hoặc giảm qua mỗi thế hệ (*Hình 4.4*).

(iii) Kiểu biến đổi tần số alen này tương quan nghịch với kích thước quần thể (đối với các nhà toán học thì nó tỷ lệ nghịch với giá trị trung bình của hàm điều hòa). Mẫu càng lớn thì kết quả càng gần giá trị trung bình (*Hình 4.4*).



Hình 4.4. **Sự biến đổi tần số alen của các alen trung tính trong 5 quần thể gồm 20 (a) và 200 (b) cá thể lưỡng bội theo máy tính. Trục tung thể hiện số bản sao của một alen có tần số khởi đầu là 0,5. Mỗi đồ thị cho thấy sự biến đổi tần số alen (số bản sao alen) ở một quần thể sau nhiều thế hệ. Ở 4 trong số 5 quần thể gồm 20 cá thể, alen đó được cố định hoặc mất đi (p = 1 hoặc = 0) trong vòng 50 thế hệ.**

Hiện tượng lạc dòng di truyền quan trọng trong ba trường hợp: (i) để loại bỏ hoặc tăng cường một alen rất hiếm, đặc biệt đối với những đột biến mới có lợi trước khi đột biến đó được củng cố, vì hiện tượng lạc dòng di truyền có hiệu ứng truyền các alen hiếm lớn hơn nhiều so với chọn lọc. Nó quan trọng trong việc làm tăng tần số các đột biến lặn mới xuất hiện tới mức có đủ tần số đồng hợp tử để quá trình chọn lọc trở nên có hiệu lực; (ii) lạc dòng di truyền chịu trách nhiệm biến đổi tần số các đột biến trung tính. Xác suất để cố định một đột biến trung tính mới xuất hiện trong quần thể bằng $1/2N_e$, với N_e là kích thước hiệu quả (để sinh sản) của quần thể (giả sử có nhiều đột biến trung tính, mỗi đột biến có một bản sao trong số $2N_e$ bản sao các alen, thì xác suất để mỗi đột biến được cố định là bằng nhau và bằng $1/2N_e$); (iii) lạc

dòng di truyền trong các quần thể nhỏ có thể sinh ra tần số alen không đặc trưng, khó bắt gặp trong các quần thể lớn. Hiện tượng đó được gọi là **hiệu ứng "kẻ sáng lập"** khi từ một nhóm nhỏ cá thể hình thành một tập đoàn mới. Ví dụ về hiện tượng này có thể nêu về trường hợp những người Amish; đó là nhóm người ở Mỹ liên kết với nhau theo tôn giáo, được hình thành từ một số ít gia đình nhập cư. Lạc dòng di truyền còn được gọi là **hiện tượng thắt cổ chai** khi một quần thể bị giảm mạnh về số lượng (do bệnh tật hoặc thiếu ăn), rồi sau đó số cá thể còn lại trở thành bố, mẹ của một quần thể lớn. Lạc dòng di truyền có thể làm cho những quần thể nhỏ bị cách ly trở nên rất khác biệt với dạng chuẩn của loài và có thể có ý nghĩa quan trọng trong việc sinh ra những chủng hoặc những loài mới. Hiện tượng nội phối có thể là một trường hợp đặc biệt của hiệu ứng người sáng lập khi mà quần thể hiệu quả chỉ là một nhóm nhỏ của các cá thể có quan hệ họ hàng, hay có thể chỉ là một cá thể lưỡng tính duy nhất.

Chọn lọc phụ thuộc tần số

Một alen có thể có ưu thế nào đó ngay khi còn hiếm, đặc biệt khi ở trạng thái dị hợp tử. Nếu một loài chim sẻ sống ở đảo, nơi có một số loài cây sinh ra các hạt có kích thước khác nhau thì những con chim có mỏ lớn có thể ăn những hạt kích thước lớn, còn những con chim mỏ nhỏ ăn những hạt kích thước nhỏ, và giữa chúng ít có sự cạnh tranh. Nếu cả những con có mỏ lớn và nhỏ cùng ăn một loại hạt thì sự cạnh tranh sẽ mạnh hơn, và chọn lọc sẽ làm giảm tần số hai kích thước mỏ đó. Ví dụ, ở loài *Geospiiza fortis*, một loài chim sẻ Darwin trên đảo Daphne Major của quần đảo Galapagos, kích thước mỏ trung bình giảm sau những năm nhiều mưa khi có nhiều hạt có kích thước nhỏ, và tăng lên sau những năm khô hạn khi chỉ có những cây cho quả lớn sống sót được.

Bệnh tật là một yếu tố chọn lọc khác góp phần duy trì biến dị trong quần thể vật chủ của nó. Các chủng sinh vật gây bệnh thích nghi với những kiểu gen vật chủ nhất định, và những chủng thành công nhất là những chủng có thể lan truyền nhanh chóng ở kiểu gen vật chủ phổ biến nhất. Chúng góp phần làm giảm giá trị chọn lọc và tần số của kiểu gen đó. Đây là một vấn đề nghiêm trọng trong nông nghiệp nếu những vùng rộng lớn chỉ trồng một giống của một loài. Trên cơ sở đó, việc chỉ đơn giản trồng lẫn ba giống lúa mạch, mỗi giống mẫn cảm với một chủng nấm gây bệnh khác nhau thì người ta đã không cần phải phun thuốc diệt nấm, vì không chủng nấm nào có thể phá hoại được toàn bộ cánh đồng. Trong một quần thể đa dạng di truyền, không một chủng sinh vật gây bệnh nào có thể lan truyền toàn diện, và đó là trạng thái thường gặp trong các quần thể tự nhiên.

Chọn lọc cân bằng – ưu thế dị hợp tử

Nếu cả hai thể đồng hợp tử cùng kém thích nghi so với thể dị hợp tử thì cả hai alen sẽ cùng được duy trì trong quần thể (ví dụ, alen bình thường và alen gây bệnh hồng cầu hình liềm). Alen quy định β-hemoglobin, Hb^S, mã hoá chuỗi polypeptid có valin ở vị trí thứ ba thế chỗ cho acid glutamic ở chuỗi polypeptid bình thường. Đồng hợp tử về alen này gây bệnh hồng cầu hình liềm, một bệnh nghiêm trọng và thường gây chết. Nhưng alen này lại phổ biến ở những nơi bệnh sốt rét lưu hành, vì dị hợp tử Hb^AHb^S chống chịu được với bệnh sốt rét. Ở nhiều vùng của châu Phi, tần số alen Hb^S đạt đến 0,2. Với tần số này thì tần số đồng hợp tử lặn là $0,2^2 = 0,04 = 4\%$; trong số này, 4/5 số cá thể có thể chết vì bệnh hồng cầu hình liềm. Đồng hợp tử trội Hb^AHb^A cũng mẫn cảm với bệnh sốt rét, chỉ có dị hợp tử Hb^AHb^S là có khả năng kháng bệnh cao hơn. Nếu tần số $Hb^S = q = 0,2$, và

giá trị chọn lọc của những người mắc bệnh hồng cầu hình liềm là 0,2 (vì 4 trong số 5 người mắc bệnh không sinh sản), vậy có thể tính được kết quả chọn lọc kiểu gen Hb^AHb^A do bệnh sốt rét như sau: tần số ở trạng thái cân bằng của $Hb^A = p = s/(s + t)$, với s là hệ số chọn lọc đối với Hb^AHb^A và t là hệ số chọn lọc đối với Hb^SHb^S. Tỷ lệ alen là nghịch đảo hệ số chọn lọc:

Kiểu gen	Hb^AHb^A	Hb^AHb^S	Hb^SHb^S
Giá trị chọn lọc	0,8	1	0,2

Trong thực tế, giá trị chọn lọc thực sự phụ thuộc vào chỉ số mắc sốt rét và sự chăm sóc y tế.

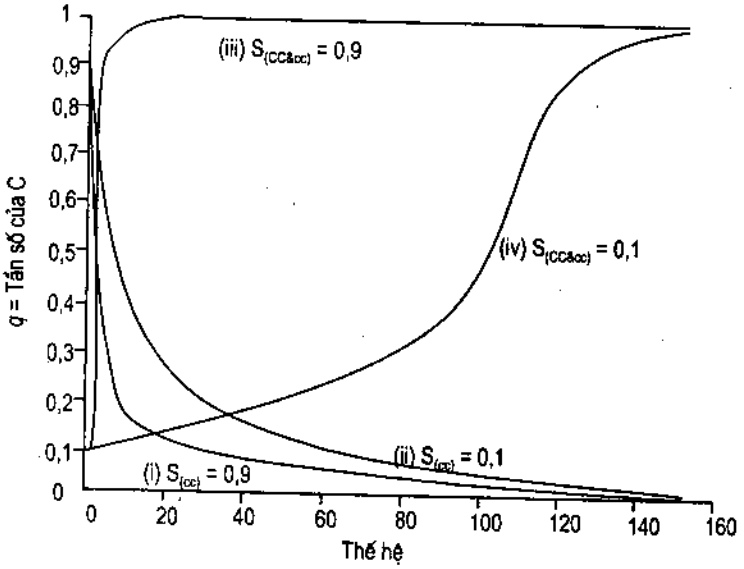
IV.E THUYẾT DARWIN MỚI: CHỌN LỌC TÁC ĐỘNG ĐẾN CÁC ALEN

Mở đầu

Các nguyên lý di truyền và đột biến được phát hiện đã bổ sung cho thuyết tiến hóa của Darwin. Tiến hoá có thể coi là sự biến đổi tần số alen và tần số gen trong quần thể qua thời gian; và chọn lọc có thể được coi là sự tác động đối với các alen và các gen cụ thể thông qua hiệu ứng trung bình trên toàn bộ cá thể mang các alen đó trong quần thể. Cách tiếp cận này đối với quá trình chọn lọc được gọi là thuyết Tiến hoá tổng hợp, hay thuyết Darwin mới.

Đơn vị chọn lọc

Người ta thường coi cá thể là đơn vị chọn lọc, vì cá thể có thể sống sót hoặc chết, có thể sinh sản nhiều hoặc ít so với cá thể khác. Điều đó chỉ đúng trong vòng một thế hệ, nhưng không có sự liên tục về phương diện cá thể trong các quần thể sinh sản hữu tính. Cá thể sinh trưởng, già đi và chết, đồng thời hệ gen của nó bị phân cắt trong quá trình phát sinh giao tử. Tuy nhiên, bản sao của các alen lại được truyền vào các tổ hợp gen mới. Tính liên tục chỉ thể hiện ở các bản sao của các alen, và thuyết Tiến hoá tổng hợp xem xét quá trình chọn lọc thông qua sự tác động đến các alen cụ thể trong mối liên quan với mức đóng góp trung bình của các alen đó cho các cá thể mang các bản sao của chúng.



Hình 4.5. **Sự biến đổi tần số alen lặn dưới tác động của các kiểu chọn lọc khác nhau (Bảng 4.3). Mỗi đồ thị đều biểu diễn sự biến đổi tần số của alen a kém thích nghi và có tần số ban đầu là 0,9. Số liệu được tính toán dựa trên việc áp dụng phương pháp chọn lọc liên tục đối với các kiểu gen ở trạng thái cân bằng Hardy - Weinberg. Hệ số chọn lọc chống lại các alen lặn là: đồ thị (i): $s_{(cc)}$ = 0,9; đồ thị (ii): $s_{(cc)}$ = 0,1, và hệ số chọn lọc chống lại các alen trội là: đồ thị (iii): $s_{(CC\&cc)}$ = 0,9; đồ thị (iv): $s_{(CC\&cc)}$ = 0,1. Với $s_{(cc)}$ = 0,9 nghĩa là 90% số cá thể có kiểu gen aa bị loại bỏ ở mỗi thế hệ.**

Mức độ phù hợp (hay Giá trị chọn lọc)

Giá trị chọn lọc có hai thành phần: sức sống và độ hữu thụ. Giá trị chọn lọc đặc trưng cho từng kiểu gen riêng biệt. Để đơn giản, cần xem xét từng gen một cách riêng biệt. Giá trị chọn lọc của một kiểu gen là tỷ lệ giá trị chọn lọc trung bình (mức độ truyền gen) của các cá thể mang kiểu gen đó so với giá trị chọn lọc của các cá thể có kiểu gen tối ưu.

Ví dụ, alen *A* mã hoá một enzym là alen trội. Đồng hợp tử *AA* và dị hợp tử *Aa* có giá trị chọn lọc là 1. Nếu enzym đó quan trọng thì các đồng hợp tử *aa* sẽ bị chết nghĩa là có giá trị chọn lọc bằng 0 và có hệ số chọn lọc $S_{(aa)} = 1 - W = 1 - 0 = 1$. Nếu enzym đó không quan trọng, nhưng có 1 trong số 100 cá thể *aa* không sinh sản được vì thiếu nó, thì giá trị chọn lọc (*Waa*) của kiểu gen *aa* là 0,99 và hệ số chọn lọc (*saa*) là 0,01. Tốc độ biến đổi tần số alen *a* trong quần thể sẽ nhanh hơn với hệ số chọn lọc cao hơn (*Hình 4.5*). Chú ý rằng chọn lọc chống lại alen trội cũng có nghĩa như chọn lọc ưu tiên cho alen lặn và ngược lại.

Sự nhiễm melanin công nghiệp

Loài bướm *B. betularia* là ví dụ kinh điển về chọn lọc tự nhiên. Chúng là loài bướm đêm và dạng điển hình là màu xám có đốm, phù hợp với việc ngụy trang trên các thân cây to có địa y bao phủ, nơi chúng trú nghỉ vào ban ngày. Năm 1849, người ta đã thu được dạng màu đen ở Manchester, Anh. Các nhà nghiên cứu đã rất vui mừng khi thu thập được dạng rất hiếm này. Dạng này được gọi là dạng carbonia (đen) và do alen trội *C* quy định, vì thế dạng điển hình có kiểu gen *cc*. Dạng bướm đen dần trở nên phổ biến, đến khoảng năm 1895, có khoảng 98% quần thể bướm có màu đen. Sự ô nhiễm, mà chủ yếu là mưa acid, từ các vùng công nghiệp đã giết chết các cây địa y trên các cây to, làm trơ các thân cây và khói từ các nhà máy nhuộm các thân cây trở nên đen hơn. Lúc này, loài bướm có màu xám điển hình không còn ngụy trang tốt bằng loài bướm đen. Những quan sát trực tiếp và thí nghiệm thả những con bướm được đánh dấu đã củng cố cho giả thuyết là: ở những vùng bị nhiễm công nghiệp, loài chim ăn thịt đã ăn thịt loại bướm màu xám điển hình nhiều hơn loại có màu đen. Khi hai loại này cùng được thả ở vùng nông thôn không bị ô nhiễm thì tình huống xảy ra ngược lại: dạng màu đen bị ăn thịt nhiều hơn.

Trong một thí nghiệm tiến hành ở Birmingham, Anh, trong số 447 bướm đen thả ra, đã bắt lại được 123 con (tỷ lệ 0,275), và trong số 137 bướm xám đã bắt được lại 18 con (tỷ lệ 0,131). Nếu dạng màu đen có giá trị chọn lọc được coi là 1 (= 0,275/0,275), thì dạng điển hình có giá trị chọn lọc là 0,131/0,275 = 0,476 (với giả thiết là sự di nhập như nhau ở hai loại bướm và hiệu suất bắt lại số bướm thí nghiệm như nhau). Một thí nghiệm tương tự cũng được tiến hành ở vùng nông thôn Doset, Anh. Trong số 473 bướm đen thả ra, đã bắt lại được 30 (tỷ lệ 0,063), và trong số 496 bướm xám điển hình đã bắt được lại 62 (tỷ lệ 0,125). Ở đây, dạng điển hình được coi là có giá trị chọn lọc bằng 1 (= 0,125/0,125), và dạng màu đen có giá trị chọn lọc bằng 0,063/0,125 = 0,504. Giá trị chọn lọc đã đảo ngược giữa hai vùng ô nhiễm và không ô nhiễm.

Những con số nêu trên không phải là tuyệt đối chính xác vì lũ bướm có thể đã đẻ trứng vào những đêm trước, trong trường hợp đó, sự chết sau sinh sản là không liên quan đến thí nghiệm. Trong một thí nghiệm khác, Bishop đã tính toán giá trị chọn lọc của loại bướm xám điển hình so với loại bướm đen dựa trên tốc độ đẻ trứng và độ sống sót của bướm trưởng

thành, giá trị đó là 0,8 ở Liverpool và 1,8 ở vùng nông thôn Wales, cách Liverpool 50km. Kết quả này vẫn còn chưa tính đến chọn lọc tác động ở các giai đoạn khác nhau trong vòng đời con bướm; ngoài ra, đã có bằng chứng cho thấy có sự ảnh hưởng đến độ sống sót ở giai đoạn nhộng. Tần số các kiểu gen ở mọi giai đoạn trong vòng đời con bướm phải được xác định qua nhiều thế hệ trước khi tính toán. Có thể tính được hiệu ứng của chọn lọc qua thời gian nếu biết được hệ số chọn lọc (*Hình 4.5*).

Chọn lọc chống lại các alen lặn

Các alen lặn chỉ chịu tác động của chọn lọc khi ở trạng thái đồng hợp tử, ví dụ *cc*. Khi tần số alen *c* (ký hiệu *q*) giảm thì giá trị q^2 rất nhỏ và phần lớn alen tồn tại ở trạng thái dị hợp tử (tần số $2pq$) không được biểu hiện và vì thế không bị chọn lọc tác động. Vì lý do đó mà chọn lọc chống lại các alen lặn rất ít hiệu quả cho dù chúng có giá trị chọn lọc cao hay thấp khi ở trạng thái đồng hợp tử (*Bảng 4.3a*). Alen đó vẫn có xu hướng kiên định với tần số thấp kể cả khi chọn lọc chống lại nó với cường độ mạnh.

Chọn lọc chống lại các alen trội

Alen màu đen trên đây là alen trội, vì thế nó chịu tác động của chọn lọc cả ở trạng thái đồng hợp tử *CC* và dị hợp tử *Cc*. Chọn lọc chống lại alen màu đen có xu hướng nhằm loại bỏ toàn bộ alen *C* vì chúng đều biểu hiện ra kiểu hình, thậm chí khi có tần số thấp (*Bảng 4.3b*).

Chú ý rằng, chọn lọc chống lại alen trội không dẫn đến việc thay thế ngay nó bằng alen lặn, vì alen lặn chủ yếu tồn tại ở trạng thái dị hợp tử và cũng bị chọn lọc đào thải. Chỉ khi tần số alen lặn đạt giá trị khoảng 0,3, nghĩa là tần số đồng hợp tử q^2 khoảng 0,09 thì tần số của nó mới bắt đầu tăng khá nhanh (*Hình 4.5*). Thậm chí khi đó hơn 2/3 số alen lặn vẫn ở trạng thái dị hợp tử ($2pq = 0,42$).

Bảng 4.3. **Sự biến đổi tần số alen sau một thế hệ do chọn lọc chống lại các kiểu gen khác nhau**

	Kiểu chọn lọc	Giá trị chọn lọc của các kiểu gen			dq
		CC	*Cc*	*cc*	
a	Trội hoàn toàn Chọn lọc chống lại *c* (ở đồng hợp tử *cc*)	1	1	$(1-s)$	$-sq^2[(1-q)/(1-sq^2)]$
b	Trội hoàn toàn Chọn lọc chống lại *C* (ở *CC* và *Cc*)	$1-s$	$1-s$	1	$sq^2(1-q)/[1-s(1-q^2)]$
c	Trội không hoàn toàn Chọn lọc chống lại *c* (cả ở *cc* và *Cc*)	1	$1-hs$	$1-s$	$-hsq(1-q)/[1-sq(2hp+q)]$

Ở đây, q là tần số alen lặn; dq là mức độ biến đổi tần số alen, s là hệ số chọn lọc; và h là mức độ trội, $h = 1$ cho trường hợp trội hoàn toàn và bằng 0 khi hai alen đồng trội; giá trị dq được tính bằng cách nhân tần số hợp tử của mỗi kiểu gen với giá trị chọn lọc tương ứng để có được tần số alen mới rồi trừ cho tần số alen ở quần thể ban đầu. Có thể tham khảo cách tính này ở các sách giáo khoa di truyền học.

Chọn lọc chống lại các alen cộng gộp

Trong trường hợp không có hiện tượng trội thì hiệu ứng của hai alen là cộng gộp, vì thế dị hợp tử có giá trị chọn lọc nằm giữa hai đồng hợp tử. Nếu là hiện tượng trội không hoàn toàn, thì chọn lọc chống lại dị hợp tử (ký hiệu *h*) sẽ nằm đâu đó giữa hai đồng hợp tử. Trong trường hợp đó, tốc độ biến đổi tần số alen do chọn lọc sẽ là trung gian giữa hiệu ứng chọn lọc chống lại các alen trội và lặn hoàn toàn.

Chọn lọc ưu tiên dị hợp tử

Hiện tượng siêu trội, hay ưu thế dị hợp tử, xảy ra khi dị hợp tử thích nghi hơn cả hai đồng hợp tử. Ví dụ kinh điển cho trường hợp này là các thể mang bệnh hồng cầu hình liềm. Những cá thể dị hợp tử có khả năng kháng bệnh sốt rét tốt hơn những người là đồng hợp tử mang các alen bình thường và không bị bệnh hồng cầu hình liềm, cũng như đối với những người đồng hợp tử lặn về alen này. Trong thực tế, alen này tồn tại chủ yếu ở trạng thái dị hợp tử. Hiện tượng này sinh ra ưu thế chọn lọc ưu tiên dị hợp tử và duy trì hiện tượng đa hình trong quần thể. Cả hai alen cùng tồn tại trong quần thể với tần số của trạng thái cân bằng được xác định bằng giá trị chọn lọc của hai đồng hợp tử tương ứng. Chọn lọc dị hợp tử có thể là một nhân tố quan trọng trong việc duy trì biến dị di truyền trong quần thể.

Chọn lọc chống lại các thể dị hợp tử

Dị hợp tử không có ưu thế chọn lọc khi nó kém thích nghi so với các đồng hợp tử. Các alen hiếm tồn tại chủ yếu ở trạng thái dị hợp tử sẽ có tốc độ đào thải cao. Nó sẽ bị đào thải hoàn toàn nếu xu thế chọn lọc này kiên định. Sự đào thải phụ thuộc vào tần số alen nhiều hơn là giá trị thích nghi của các đồng hợp tử. Tình huống này xuất hiện khi hai quần thể liền kề bị cách ly và trở thành đồng hợp tử về các alen khác nhau, rồi sau đó tiếp xúc trở lại ở vùng ranh giới phân bố của hai quần thể. Nếu một số thành viên của hai nhóm đồng hợp tử này không giao phối được với nhau (nghĩa là đã có cơ chế cách ly tiền giao phối), thì hai quần thể có khả năng trở thành hai loài.

IV.F SỰ BIẾN ĐỔI NHIỄM SẮC THỂ TRONG QUÁ TRÌNH TIẾN HÓA

Kiểu nhân

Thuật ngữ **kiểu nhân** dùng để mô tả số lượng, kích thước, và vị trí tâm động/tỷ lệ độ dài các cánh của nhiễm sắc thể. **Các nhiễm sắc thể thường** là các nhiễm sắc thể có ở mọi cá thể, cả đực và cái. **Nhiễm sắc thể tâm giữa** có tâm động nằm gần chính giữa và hai cánh có độ dài gần bằng nhau. **Nhiễm sắc thể tâm cận giữa** có hai cánh có độ dài không bằng nhau. **Nhiễm sắc thể tâm lệch** có một cánh dài và một cánh rất ngắn. **Nhiễm sắc thể tâm mút** có tâm động nằm kề một **đầu mút**. Đầu mút là đoạn trình tự đặc biệt, liên kết với protein để bảo vệ đầu của nhiễm sắc thể khỏi bị phân hủy; nó tương tác với màng nhân và bộ khung tế bào trong quá trình giảm phân.

Số nhiễm sắc thể có trong một bộ đơn nhiễm sắc thể ở giao tử được gọi là số lượng đơn bội và được ký hiệu là *n*, vì thế thể **lưỡng bội** (hợp tử do hai giao tử đơn bội kết hợp) có $2n$ nhiễm sắc thể. Nói cách khác, các thể lưỡng bội có *n* cặp nhiễm sắc thể. Ở một số loài, hai

giới tính có số lượng nhiễm sắc thể giới tính khác nhau, vì thế con đực và con cái có thể có số lượng lưỡng bội nhiễm sắc thể khác nhau. **Các nhiễm sắc thể giới tính** có thể được ký hiệu riêng (ví dụ, một loài có $n = 8$, con cái có 2 nhiễm sắc thể X, con đực có 1 nhiễm sắc thể X và không có Y, thì kiểu nhân con đực là $2n = 15XO$ và con cái là 16XX). Kiểu nhân là đặc điểm cố định của loài, tuy nhiên có thể có các biến dạng nhiễm sắc thể mang các thông tin tương đồng tồn tại trong loài. Đó là hiện tượng **đa hình nhiễm sắc thể**.

Kiểu nhân có thể rất ổn định trong quá trình tiến hóa. Hầu hết các loài chuồn chuồn có 13 nhiễm sắc thể, rắn và chim đều có 18 nhiễm sắc thể lớn và một số nhiễm sắc thể nhỏ. Các số liệu này cho thấy sự bảo tồn kiểu nhân trải qua nhiều triệu năm tiến hoá. Mặt khác, những loài có quan hệ họ hàng gần có thể có những biến đổi rất nhanh. Loài *Muntiacus muntjac vaginalis* ở Ấn Độ, giới đực 2n = 7, còn giới cái có 2n = 6; trong khi loài *Muntiacus muntjac reevesi* ở Trung Quốc có 2n = 46. Các chủng Dế (*Grylotalpa grylotalpa*) khác nhau ở Châu Âu có số lưỡng bội nhiễm sắc thể từ 12 đến 23. Biến dị này có thể liên quan đến sự hình thành loài.

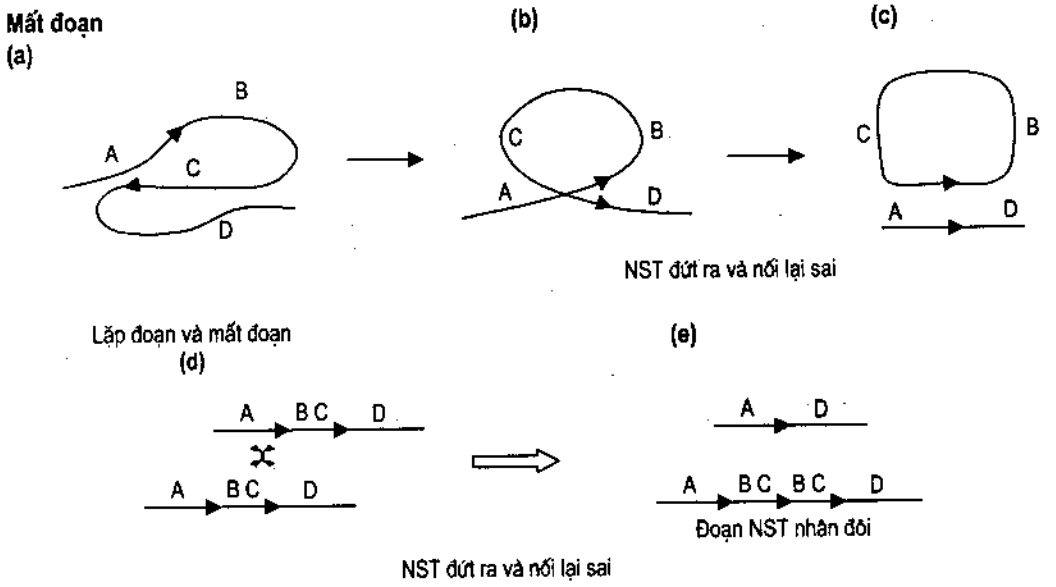
Hiện tượng đa hình nhiễm sắc thể và tính hữu thụ

Sự sinh sản hữu tính cần quá trình giảm phân để phân ly thông tin di truyền trong ADN của các nhiễm sắc thể sao cho mỗi giao tử có một bản sao của mỗi gen và để mọi hợp tử đều có hai bản sao của mỗi gen đó (trừ gen trên nhiễm sắc thể X). Nếu tế bào là dị hợp tử do tái cấu trúc nhiễm sắc thể thì các gen tương đồng không nằm ở các vị trí tương ứng và sự phân ly có thể bị phá hủy, sinh ra các giao tử không cân bằng về gen; có giao tử mất đi một phần thông tin, trong khi giao tử khác lại mang gấp đôi số thông tin đó. Vấn đề sẽ phức tạp thêm nếu có trao đổi chéo ở các nhiễm sắc thể dị hình (mang tái cấu trúc) như vậy. Đó thực sự là vấn đề ở các thể dị hợp tử mang tái cấu trúc nhiễm sắc thể, làm cho các tái cấu trúc mới khó được sản sinh thêm trong quần thể (hầu hết các yếu tố di truyền hiếm thường ở trạng thái dị hợp tử). Các hiện tượng đa hình nhiễm sắc thể tìm được ở các quần thể, thông thường, hoặc đó là biểu hiện của trạng thái chuyển giao trong quá trình thay thế kiểu nhiễm sắc thể này bằng kiểu nhiễm sắc thể mới khác, hoặc là biểu hiện của cơ chế di truyền đặc biệt nào đó đang hoạt động để tạo ra các giao tử cân bằng, hoặc để giảm thiểu sự mất tính hữu thụ. Sự khác biệt nhiễm sắc thể giữa các chủng có thể hoạt động như cơ chế cách ly sau giao phối gây nên tính bất thụ ở con lai.

Các cơ chế tái cấu trúc nhiễm sắc thể

Sự tái cấu trúc nhiễm sắc thể cần ít nhất hai đứt gãy (tạo ra 4 đầu tự do) trên phân tử ADN (nhiễm sắc thể), sau đó chúng nối lại với nhau. Sự đứt-nối này có thể xảy ra do tái tổ hợp giữa các trình tự ADN lặp lại hoặc phân tán trong hệ gen. Các yếu tố di truyền vận động có thể trực tiếp xúc tiến hiện tượng tái cấu trúc như vậy của nhiễm sắc thể nhờ cơ chế vận động của nó. Việc sửa chữa sai các đầu tự do của chuỗi xoắn kép bị đứt gãy do bức xạ ion hoá (hoặc các tác nhân đột biến khác) là nguồn tái tổ hợp khác đã được chứng minh có ở các tế bào nuôi cấy. Tất cả các nhiễm sắc thể mạch thẳng đều có đầu mút là đoạn trình tự lặp lại đặc biệt liên kết với protein để bảo vệ các đầu của nhiễm sắc thể khỏi bị phân hủy. Các nhiễm sắc thể đứt gãy không được nối lại hoặc không được bổ sung cấu trúc đầu mút thường không bền vững và không hoạt động được. Hơn nữa, chúng còn thúc đẩy sự đứt gãy và tái cấu trúc nhiễm sắc thể

ở các lần phân bào tiếp theo. Người ta phân biệt tái cấu trúc phần đầu mút và tái cấu trúc phần bên trong (không liên quan đến đầu mút) của nhiễm sắc thể. Hiện tượng tái cấu trúc nhiễm sắc thể có thể chỉ liên quan đến một nhiễm sắc thể duy nhất, nhưng cũng có thể liên quan đến sự chuyển đoạn hay trao đổi đoạn giữa các nhiễm sắc thể không tương đồng (rất khác nhau về thông tin di truyền).



Hình 4.6. **Mất đoạn BC dưới dạng vòng tròn khép kín (a - c) có thể được tạo ra do đứt nối hoặc tái tổ hợp giữa các đoạn trình tự lặp lại trên cùng một NST, hoặc do các yếu tố di truyền vận động; vùng mất đoạn/lặp đoạn BC (d – e) có thể do đứt nối hoặc tái tổ hợp không theo hàng trên các NST tương đồng hoặc trên các nhiễm sắc tử chị em.**

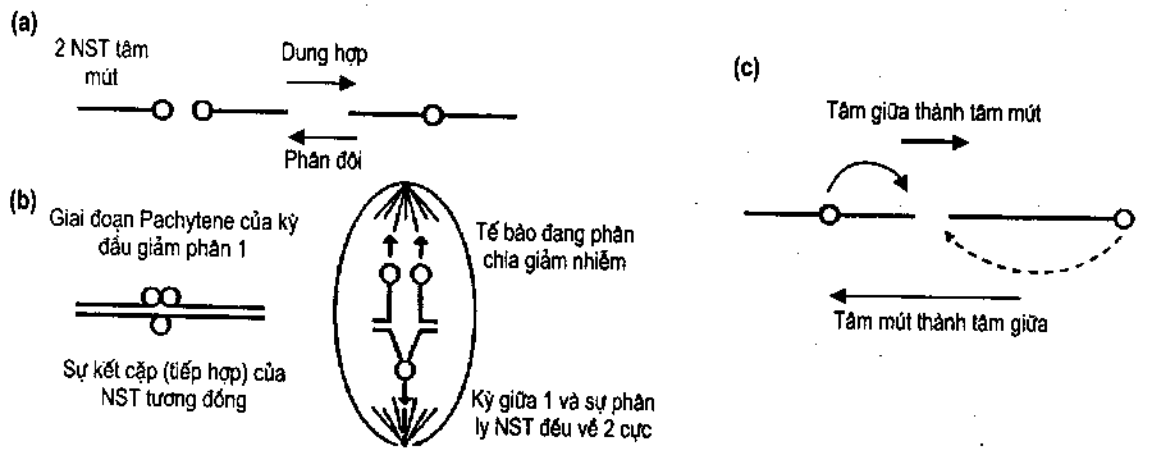
Mất đoạn

Trong trường hợp mất đoạn chứa đầu mút gồm một đứt gãy, sau đó đoạn đứt mang tâm động tạo đầu mút mới, thì **đoạn không tâm** (không chứa tâm động) sẽ bị mất trong quá trình phân bào. Các mất đoạn vùng bên trong của nhiễm sắc thể, sau đó là sự nối lại các đầu tự do của các đoạn kề nhau. Sự nối lại có thể diễn ra do tái tổ hợp các đoạn trình tự lặp lại theo cùng một hướng trên cùng một nhiễm sắc thể, loại bỏ đi một đoạn dưới dạng vòng tròn kín (*Hình 4.6: a - c*). Ngoài ra, sự tái tổ hợp không thẳng hàng giữa các đoạn trên các nhiễm sắc thể tương đồng có thể tạo ra cặp lặp đoạn - mất đoạn (*Hình 4.6: d* và *e*). Lặp đoạn cũng có thể được sinh ra do trao đổi chéo trong giảm phân ở các vùng nhiễm sắc thể dị hợp tử về các kiểu tái cấu trúc khác, như chuyển đoạn hoặc đảo đoạn (xem phần sau).

Dị hợp tử mang chuyển đoạn thường là đơn bội một phần cho đến khi vùng đó được lặp đoạn trở lại. Ở ruồi giấm *Drosophila melanogaster* có hai locut gen gây chết trội ngay khi chỉ có một bản sao là locut 11 và locut 41 (sinh ra ruồi có kích thước rất bé). Nhiều gen có một bản sao thường không đủ để ruồi phát triển bình thường; ngoài ra, kết hợp với việc mất thêm khoảng 3% hệ gen thì thường gây chết. Có lẽ, đó là hiệu ứng phối hợp của việc mất đi một bản sao của nhiều gen. Có lẽ hiện tượng mất đoạn ít nguy hại ở những loài đa bội. Mất đoạn dị hợp tử, đặc biệt là những mất đoạn lớn, có thể ngăn cản sự tiếp hợp của các nhiễm sắc thể tương đồng trong giảm phân, gây nên sự không phân ly và tạo nên độ hữu thụ thấp.

Lặp đoạn

Có lẽ từ trước khi sự sống dạng tế bào xuất hiện, các lặp đoạn là nguồn nguyên liệu vật chất di truyền mới chủ yếu, được hình thành do sự tiến hóa phân ly và phân tách của hai bản sao vật chất di truyền. Kết quả giải trình tự toàn bộ hệ gen nấm men *Saccharomyces cereviseae* cho thấy có 53 nhóm gen lặp đoạn trên các nhiễm sắc thể không tương đồng. Trong một số trường hợp, khi cần nhiều bản sao của một gen (như gen rARN), thì các bản sao phân bố liên tiếp được hình thành một cách hài hoà trong loài, tất cả chúng đều có cùng một trình tự. Sự tái tổ hợp không theo hàng (giữa các nhiễm sắc tử chị em) giữa các đoạn trình tự lặp lại phân tán trên các nhiễm sắc thể tương đồng có lẽ là nguồn gốc của các lặp đoạn kế tiếp nhau trên một nhiễm sắc thể (*Hình 4.6d* và *e*). Đôi khi, một chuyển đoạn nhỏ cũng tồn tại được ở vị trí mới dưới dạng một lặp đoạn. Khi tiến hành chọn lọc theo hướng tăng số bản sao của một gen, các lặp đoạn nhiều bản sao của gen đó thường xuất hiện dưới kính hiển vi dưới dạng các vùng nhuộm màu không đồng nhất, chứa nhiều đoạn lặp lại kế tiếp. Một ví dụ là thí nghiệm chọn lọc giống Rệp vừng bằng thuốc trừ sâu nhằm làm tăng số gen mã hóa enzym esterase có tác dụng phá huỷ thuốc giúp loài côn trùng này trở nên kháng thuốc; hoặc bằng liệu pháp hoá học đối với các tế bào ung thư để khuếch đại các gen mã hóa protein làm nhiệm vụ bơm các chất độc ra khỏi tế bào. Các lặp đoạn vùng tâm động không sống sót được vì hai tâm động có thể định hướng ngược nhau trong quá trình phân bào, bẻ gãy nhiễm sắc thể. Các lặp đoạn thường ít nguy hiểm hơn so với mất đoạn.

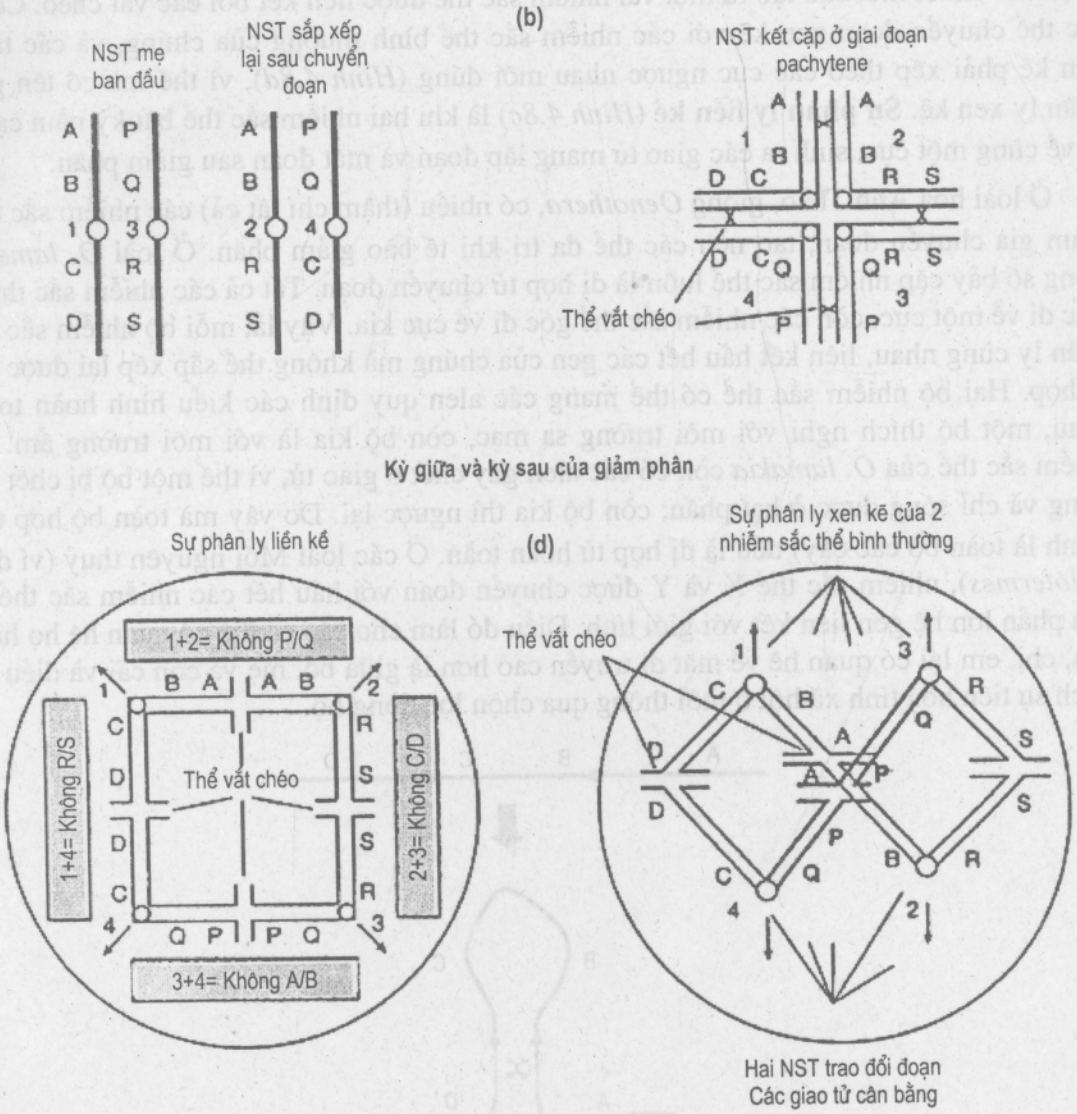


Hình 4.7. **(a) Sự chuyển đổi diễn ra giữa hai NST tâm mút và một NST tâm giữa do sự kết hợp hoặc phân cắt tâm động (vòng tròn biểu diễn tâm động). (b) Giảm phân ở thể dị hợp tử sẽ sinh ra các giao tử cân bằng nếu hai NST tâm mút cùng đi về cực đối diện với NST tâm giữa. (c) Sự vận động của tâm động trong quá trình đảo đoạn mang tâm có thể làm thay đổi hình dạng NST từ tâm mút thành tâm giữa và ngược lại.**

Sự kết hợp tâm động và phân đôi tâm động

Hai nhiễm sắc thể tâm lệch hoặc tâm mút có cánh ngắn không mang vật chất di truyền quan trọng có thể kết hợp với nhau ở tâm động, sinh ra một nhiễm sắc thể tâm giữa hoặc tâm cận giữa, làm giảm số lượng nhiễm sắc thể (*Hình 4.7a*). Hiện tượng đó còn được gọi là **sự kết hợp nhiễm sắc thể kiểu Robertson** (hay **chuyển đoạn Robertson**). Hiện tượng ngược lại là **sự phân đôi tâm động**, trong đó tâm động của một nhiễm sắc thể tâm cận giữa phân cắt thành hai, sinh ra hai nhiễm sắc thể tâm mút, làm tăng số lượng nhiễm sắc thể (*Hình 4.7a*). Khi là dị hợp tử, hai nhiễm sắc thể tâm mút có thể kết đôi với chiếc tương đồng tâm giữa, tạo nên thể tam trị trong giảm phân.

Nếu hai nhiễm sắc thể tâm mút phân ly về cùng một cực, ngược hướng với chiếc tâm giữa thì sự phân ly diễn ra bình thường và các giao tử, hoặc chứa hai nhiễm sắc thể tâm mút, hoặc chứa một nhiễm sắc thể tâm giữa sẽ cân bằng về gen (*Hình 4.7b*). Nếu một nhiễm sắc thể tâm mút phân ly cùng chiếc nhiễm sắc thể tâm giữa thì sẽ sinh ra giao tử mang hai bản sao của các gen trên nhiễm sắc thể tâm mút và sẽ sinh ra thể ba nhiễm. Loại giao tử kia sẽ không có bản sao nào của chiếc nhiễm sắc thể tâm mút đó và sinh ra hợp tử là thể một nhiễm. Ở chuột và người, sự kết hợp tâm động kiểu Robertson ở trạng thái dị hợp tử biểu hiện tính bất thụ cao hơn nhiều so với tính toán theo lý thuyết, nhưng nguyên nhân còn chưa rõ. Đảo đoạn với vết đứt cạnh tâm động có thể chuyển nhiễm sắc thể tâm giữa thành tâm mút và ngược lại (*Hình 4.7c*).

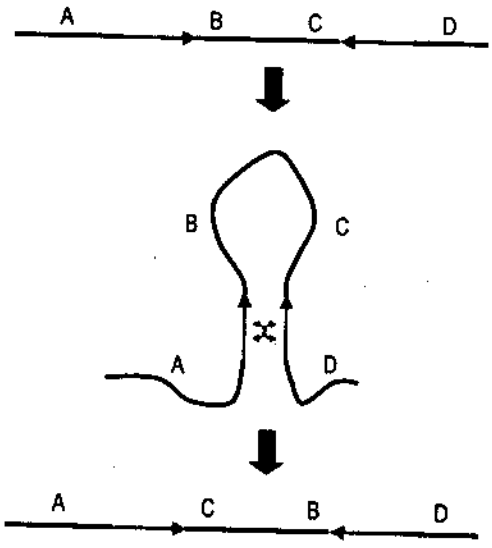


Hình 4.8. (a) Một chuyển đoạn bao gồm sự trao đổi đoạn giữa hai nhiễm sắc thể không tương đồng. (b) Hai nhiễm sắc thể bình thường và hai nhiễm sắc thể mang tái cấu trúc ghép đôi tạo ra vòng tứ trị hoặc chuỗi ở giảm phân. Các tâm động được đánh số theo trình tự trong vòng. (c) Nếu hai nhiễm sắc thể liền kề bất kỳ phân ly về cùng một cực sẽ sinh ra các giao tử không cân bằng mang một lặp đoạn và một mất đoạn. (d) Sự phân ly của các nhiễm sắc thể xen kẽ về cùng một cực sẽ sinh ra các giao tử mang một bản sao của tất cả các đoạn ABCD và PQRS.

Chuyển đoạn

Chuyển đoạn là sự chuyển đổi vật chất di truyền giữa hai nhiễm sắc thể không tương đồng. Hiện tượng này thường là sự trao đổi thuận nghịch hai đầu của nhiễm sắc thể (*Hình 4.8a*). Các thể dị hợp tử chỉ sinh ra các giao tử cân bằng gen khi cả hai nhiễm sắc thể gốc hoặc cả hai nhiễm sắc thể mang tái cấu trúc cùng được phân ly vào một giao tử. Hiện tượng tái tổ hợp làm phức tạp sự phân ly này, nhưng có một cách phân ly, đó là sự phân ly liền kề, đảm bảo sinh ra các giao tử cân bằng. Cách phân ly này dành cho các nhiễm sắc thể tâm giữa và không có trao đổi chéo giữa điểm chuyển đoạn và tâm động (*Hình 4.8b*). Các thể dị hợp tử chuyển đoạn có thể hình thành các nhiễm sắc thể dạng **vòng hoặc chuỗi đa trị** trong giảm phân. Đó là các vòng khép kín hoặc các chuỗi mở, cấu tạo từ một vài nhiễm sắc thể được liên kết bởi các vắt chéo. Các nhiễm sắc thể chuyển đoạn xen kẽ với các nhiễm sắc thể bình thường của chúng, và các tâm động xen kẽ phải xếp theo các cực ngược nhau mới đúng (*Hình 4.8d*), vì thế mà có tên gọi là sự phân ly xen kẽ. **Sự phân ly liền kề** (*Hình 4.8c*) là khi hai nhiễm sắc thể bất kỳ nằm cạnh nhau đi về cùng một cực, sinh ra các giao tử mang lặp đoạn và mất đoạn sau giảm phân.

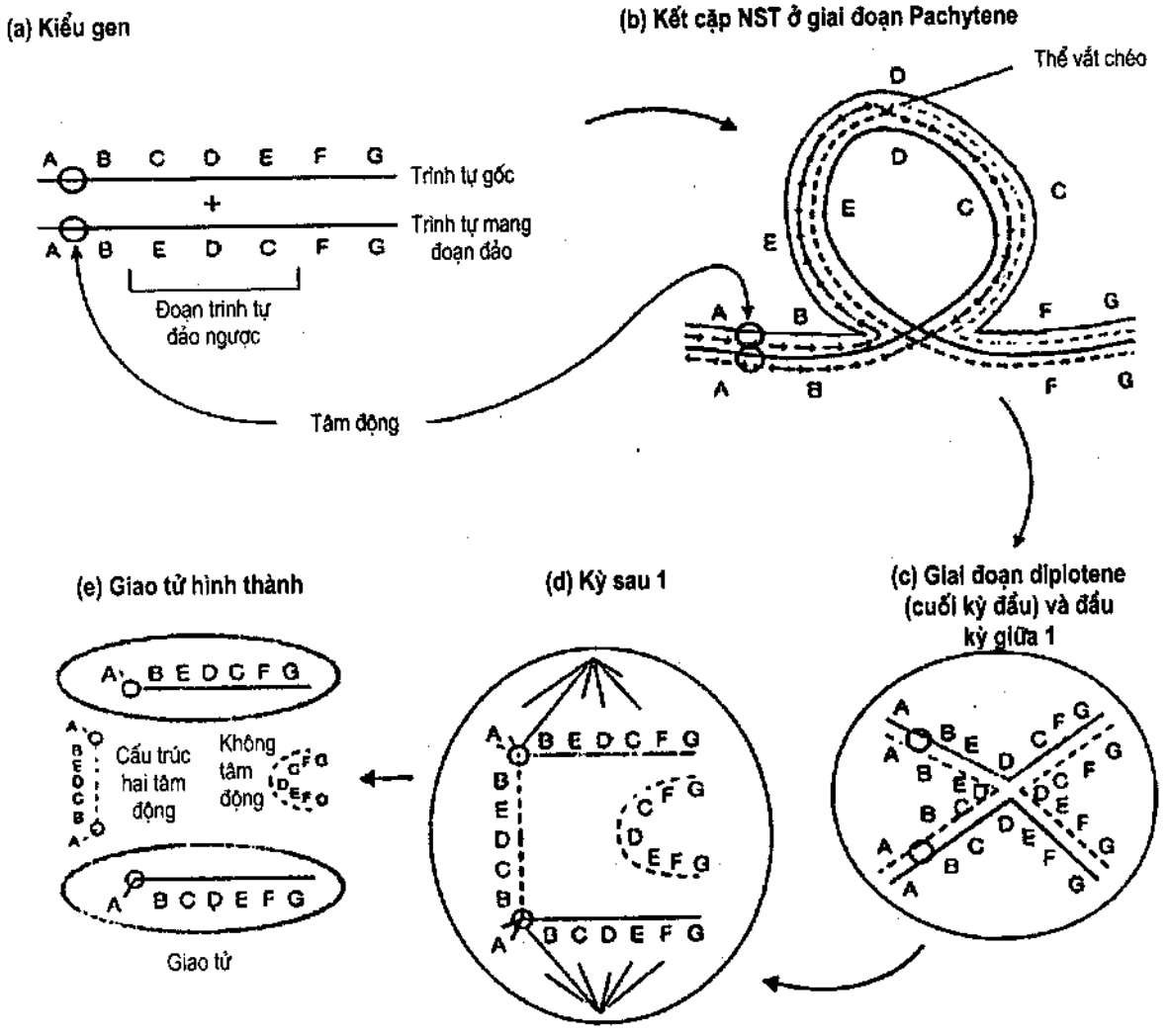
Ở loài hoa Anh Thảo, giống *Oenothera*, có nhiều (thậm chí tất cả) các nhiễm sắc thể cùng tham gia chuyển đoạn, tạo nên các thể đa trị khi tế bào giảm phân. Ở loài *O. lamakia*, sáu trong số bảy cặp nhiễm sắc thể luôn là dị hợp tử chuyển đoạn. Tất cả các nhiễm sắc thể tái cấu trúc đi về một cực, còn các nhiễm sắc thể gốc đi về cực kia. Vậy là, mỗi bộ nhiễm sắc thể luôn phân ly cùng nhau, liên kết hầu hết các gen của chúng mà không thể sắp xếp lại được bằng tái tổ hợp. Hai bộ nhiễm sắc thể có thể mang các alen quy định các kiểu hình hoàn toàn khác nhau, một bộ thích nghi với môi trường sa mạc, còn bộ kia là với môi trường ẩm. Mỗi bộ nhiễm sắc thể của *O. lamakia* còn có các alen gây chết ở giao tử, vì thế một bộ bị chết ở tế bào trứng và chỉ sống được ở hạt phấn; còn bộ kia thì ngược lại. Do vậy mà toàn bộ hợp tử (cũng chính là toàn bộ các cây) đều là dị hợp tử hoàn toàn. Ở các loài Mối nguyên thuỷ (ví dụ giống *Kalotermes*), nhiễm sắc thể X và Y được chuyển đoạn với hầu hết các nhiễm sắc thể thường nên phần lớn hệ gen liên kết với giới tính. Điều đó làm cho các cá thể có quan hệ họ hàng như anh, chị, em lại có quan hệ về mặt di truyền cao hơn là giữa bố, mẹ và con cái và điều đó kích thích sự tiến hoá tính xã hội ở mối thông qua chọn lọc dòng họ.



Hình 4.9. **Sự sản sinh đảo đoạn do tái tổ hợp giữa hai vùng trình tự lặp lại ngược chiều trong cấu trúc thòng lọng hình omêga (Ω).**

Đảo đoạn

Đảo đoạn xảy ra khi một phần của một nhiễm sắc thể bị quay ngược lại, có lẽ do tái tổ hợp qua thòng lọng có hình chữ omêga (*Hình 4.9*). Các đảo đoạn dị hợp tử có thể kết đôi ở giảm phân, tạo **thòng lọng đoạn đảo**. Tái tổ hợp ở vùng thòng lọng đoạn đảo sẽ sinh ra lặp đoạn ở nhiễm sắc thể này và mất đoạn ở nhiễm sắc thể kia, vì thế các nhiễm sắc tử tái tổ hợp không sống sót được. Kết quả chính xác còn phụ thuộc vào việc tâm động nằm trong (**đảo đoạn mang tâm**) hay nằm ngoài (**đảo đoạn không tâm**) vùng đảo đoạn. Cũng giống như các trạng thái dị hợp tử làm giảm tính hữu thụ khác, các đảo đoạn hiếm dường như bị chọn lọc đào thải. Một số loài tránh tác động có hại ấy bằng cách không kết đôi đoạn đảo. Đảo đoạn là yếu tố ức chế tái tổ hợp vì chỉ có các nhiễm sắc thể mang đoạn đảo dạng bố, mẹ không bị tái tổ hợp mới truyền được cho thế hệ sau.



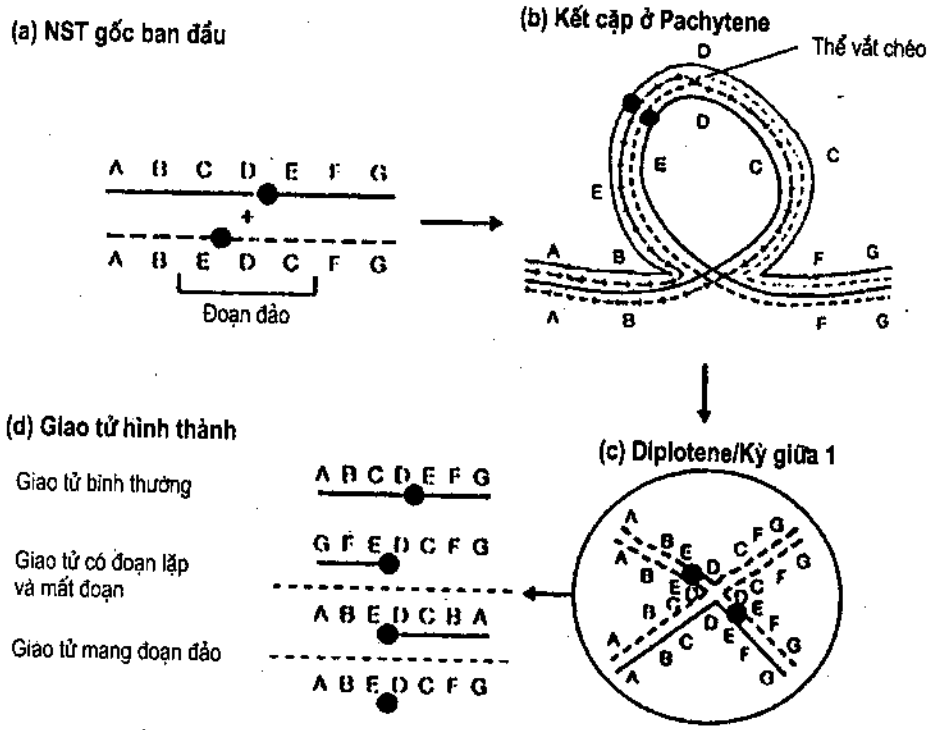
Hình 4.10. Hiệu quả do tái tổ hợp trong vùng đảo đoạn không tâm dị hợp tử ở ruồi giấm cái *Drosophila*. (a) Cặp nhiễm sắc thể đa hình: nhiễm sắc thể phía dưới mang đảo đoạn. (b) Vùng đảo đoạn kết đôi, tạo thòng lọng đảo đoạn ở giai đoạn Pachytene. Mỗi nhiễm sắc thể có hai nhiễm sắc tử. Trao đổi chéo sinh ra một nhiễm sắc tử hai tâm ABDECBA và một đoạn không tâm GFDECFG. (c) Ở giai đoạn Diplotene, các nhiễm sắc thể tương đồng ghép đôi và có thể thấy rõ vắt chéo ở vùng đảo đoạn. (d) Ở kỳ sau, nhiễm sắc tử hai tâm tạo thành cầu nối giữa hai cực và đứt ra, đoạn không tâm bị mất. (e) Một bộ nhiễm sắc tử không bị tái tổ hợp mang đầy đủ các gen trở thành giao tử.

Đảo đoạn không tâm

Đảo đoạn không tâm có các điểm đứt trên cùng một phía của tâm động (*Hình 4.10a*). Tái tổ hợp ở giảm phân trong thông lọng đoạn đảo (giai đoạn pachytene, *Hình 4.10b*) sẽ sinh ra một nhiễm sắc tử tái tổ hợp mang hai tâm động và nhiễm sắc tử kia không có tâm động (*Hình 4.10c-e*). Đoạn không tâm bị tụt lại ở kỳ sau I và bị mất; cầu nối hai tâm bị đứt, sinh ra hai nhiễm sắc thể không hoàn chỉnh. Các côn trùng hai cánh thường là đa hình về các đoạn đảo không tâm có vai trò quan trọng trong sinh thái học và tiến hoá của chúng. Ở con cái, cầu nối hai tâm giữ cho hai tâm động nằm ở giữa vùng thoi vô sắc (*Hình 4.10d*), vì thế mà ở hai cực đều là các nhiễm sắc tử không bị tái tổ hợp. Trong quá trình giảm phân ở con cái, một tế bào sẽ thành tế bào trứng, còn ba tế bào kia thành thể cực. Cơ chế này đảm bảo cho tế bào trứng sống sót được vì mang tất cả các nhiễm sắc thể không bị tái tổ hợp, còn các sản phẩm không sống sót được đều bị đào thải dưới dạng các thể cực. Ở côn trùng hai cánh, giới đực nói chung không có tái tổ hợp, vì thế nó tránh được hậu quả của hiện tượng này.

Đảo đoạn mang tâm

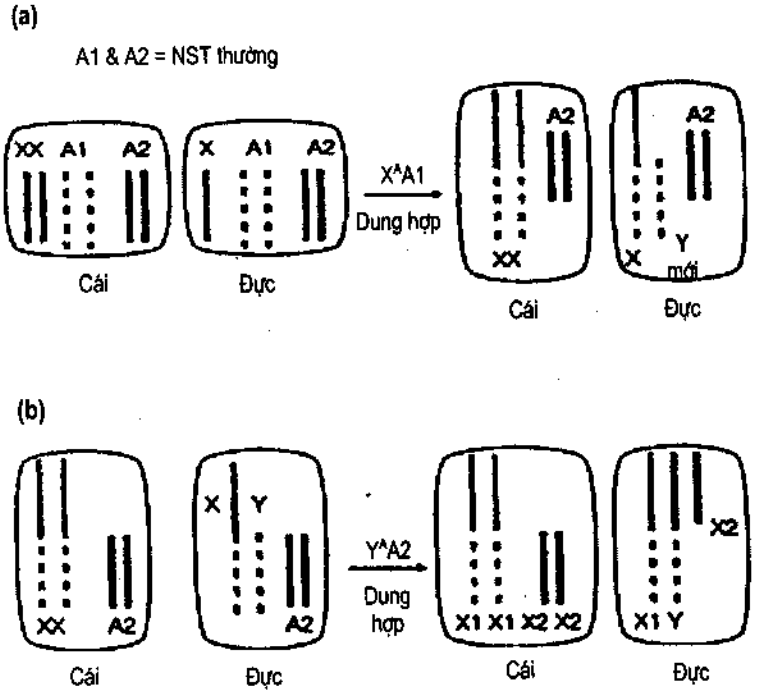
Đảo đoạn mang tâm chứa tâm động (*Hình 4.11a*). Trao đổi chéo trong thông lọng đảo đoạn làm cho các đầu của một nhiễm sắc tử thuộc mỗi nhiễm sắc thể trao đổi nhau (*Hình 4.11b* và *c*), gây lặp đoạn một đầu và mất đoạn ở đầu kia (giống như xảy ra với trao đổi chéo không tâm). Tuy nhiên, với trao đổi chéo mang tâm thì tất cả các nhiễm sắc tử đều có một tâm động (*Hình 4.11d* và *4.10d*), nên chúng có thể phân ly bình thường ở giảm phân vào các giao tử, tạo ra các giao tử không cân bằng gen ở cả hai giới tính, vì thế sinh ra các hợp tử không có sức sống. Hiện tượng đa hình các đảo đoạn mang tâm không tìm thấy ở ruồi giấm nhưng có ở các loài khác, chẳng hạn như châu chấu.



Hình 4.11. **Hiệu quả tái tổ hợp ở đảo đoạn mang tâm dị hợp tử. (a) Sự kiện này giống với đảo đoạn không tâm (*Hình 4.10*) nhưng tâm động nằm trong thông lọng đảo đoạn ở giai đoạn Pachytene. (b) Trao đổi chéo bên trong đoạn đảo sinh ra lặp đoạn và mất đoạn (c), nhưng mọi nhiễm sắc tử đều có một tâm động, vì thế các nhiễm sắc tử mang lặp đoạn và mất đoạn đều phân ly được vào các giao tử (d).**

Sự biến đổi ở nhiễm sắc thể giới tính

Loại nhiễm sắc thể chỉ có ở giới tính dị giao tử (như nhiễm sắc thể Y ở người và ruồi giấm) chỉ mang rất ít gen; ở một số loài nó bị mất đi, sinh ra con đực XO và con cái XX, đồng thời làm giảm số lượng nhiễm sắc thể. Sự kết hợp tâm động và chuyển đoạn giữa nhiễm sắc thể giới tính với các nhiễm sắc thể thường có thể làm tăng số lượng hoặc kích thước nhiễm sắc thể giới tính do nhiễm sắc thể thường chuyển đổi thành các nhiễm sắc thể giới tính mới. Những nhiễm sắc thể mới này thường bị tác động để tiếp tục duy trì trạng thái tự do như cũ, nhằm cân bằng sự phân ly của các gen tương đồng gắn với nhiễm sắc thể giới tính và duy trì trạng thái lưỡng bội. Ví dụ, khi nhiễm sắc thể X kết hợp tâm động với một nhiễm sắc thể thường thì nhiễm sắc thể mới chỉ cân bằng ở giới cái, còn ở giới đực chỉ có một nhiễm sắc thể thường nằm tự do và hoạt động giống như nhiễm sắc thể Y (*Hình 4.12a*). Nó được gọi là nhiễm sắc thể Y mới. Nếu đã có một nhiễm sắc thể Y thì nhiễm sắc thể Y mới được đánh số là Y2 (nhiễm sắc thể Y cũ là Y1). Nếu nhiễm sắc thể Y kết hợp tâm động với một nhiễm sắc thể thường (*Hình 4.12b*), giới đực sẽ có một nhiễm sắc thể thường tự do, còn giới cái có hai và được gọi là các nhiễm sắc thể X mới. Theo cách đó, số lượng nhiễm sắc thể X và Y có thể tăng lên. Chúng được đánh số lần lượt là X1, X2, X3..., và Y1, Y2, Y3...



Hình 4.12. **(a) Sự kết hợp tâm động của nhiễm sắc thể X với một nhiễm sắc thể thường A1 (tạo ra nhiễm sắc thể X^A1) làm cho chiếc nhiễm sắc thể thường kia hoạt động như một nhiễm sắc thể Y mới ở giới đực. (b) Sự kết hợp tâm động của nhiễm sắc thể Y với nhiễm sắc thể thường A2 (tạo ra Y^A2) làm cho nhiễm sắc thể A2 tự do kia hoạt động như nhiễm sắc thể X ở giới cái; như vậy, có hai nhiễm sắc thể X mới ở giới cái và một ở giới đực để cân bằng với bản sao gắn vào Y.**

Hiệu ứng tiến hóa

Các côn trùng hai cánh không có tái tổ hợp ở giới đực nên sự thích nghi đều tránh làm mất khả năng hữu thụ do chuyển đoạn mang tâm. Giới cái tránh truyền các nhiễm sắc tử tái tổ hợp mang lặp

đoạn và mất đoạn vào tế bào trứng, vì vậy mà các đoạn đảo mang tâm có thể duy trì ở côn trùng hai cánh, thậm chí là khá phổ biến. Sự ức chế tái tổ hợp giữa nhiễm sắc thể mang đoạn đảo và nhiễm sắc thể bình thường cho phép các gen bị khoá lại trong đoạn đảo cùng tiến hoá. Các gen đó có thể trở nên thích nghi với những môi trường sống khác nhau (ví dụ, môi trường nóng hoặc lạnh, độ cao khác nhau...). Có những đoạn đảo mang tâm đặc biệt ở *Drosophila* liên quan với sự xác định giới tính, đóng vai trò là nhiễm sắc thể giới tính. Nhiều quần thể có sự đa hình chuyển đoạn ổn định. Có khoảng 2000 loài *Drosophila* với số lượng đoạn đảo ước tính khoảng 350 triệu, trong số đó có khoảng 20.000 - 60.000 đoạn đảo đã được cố định và khoảng 18.000 - 28.000 đoạn đảo đang tiến hoá. Ở những loài khác nhau, tính phổ biến của các kiểu tái cấu trúc nhiễm sắc thể có khác nhau. Ví dụ, trong nhóm linh trưởng, ở các loài khỉ, kể cả người, chủ yếu là đoạn đảo không tâm; ở loài Vượn Cáo, chủ yếu là hiện tượng kết hợp tâm động, còn ở *Cercopithidae*, lại có nhiều hiện tượng phân tách tâm động.

IV.G LOÀI VÀ SỰ HÌNH THÀNH LOÀI

Theo di truyền học, loài là một tập hợp các cá thể sinh vật thực sự hoặc có tiềm năng giao phối với nhau. Các thành viên của loài cùng chia sẻ một vốn gen chung. Các thành viên của các loài khác nhau không trao đổi vật chất di truyền với nhau. Ở hầu hết các trường hợp, các loài có sự khác biệt rõ ràng (như giữa chó và mèo), tuy nhiên nhiều loài có các chủng địa lý trông khác nhau song lại có thể giao phối với nhau khi tiếp xúc. Điều đó làm cho việc xác định loài và việc bảo tồn trở nên khó khăn. Ví dụ, Vịt đực Mallard từ Châu Âu giao phối được với Vịt cái đen Canada, và Vịt Ruddy từ Châu Mỹ giao phối được với Vịt đầu trắng Châu Âu. Trong cả hai trường hợp, hiện tượng lai đã phá huỷ ranh giới các loài địa phương. Ngược lại, một số sinh vật mà con người không thấy khác biệt (như những giống Chim kích thước nhỏ có màu nâu, những động vật có vú hoặc côn trùng cỡ nhỏ) lại có thể là những loài khác nhau. Một tên loài có thể chứa đựng trong nó nhiều loài ẩn mà ta không phân biệt được. Loài Dơi phổ biến nhất ở Anh, *Pipistrellus pipistrellus*, gần đây được phát hiện là gồm hai loài, mỗi loài phát tín hiệu siêu âm khác nhau.

Loài

Phân loại học là sơ đồ xác định các loài và nhóm các loài có quan hệ họ hàng với nhau. Sơ đồ này nhằm phân nhóm các loài có chung tổ tiên. Bước thứ nhất để xây dựng sơ đồ dựa trên quan điểm về **loài hình thái**. Các thành viên của một loài có những đặc điểm cấu tạo cơ thể chung có thể phân biệt được với các thành viên của loài khác. Đây là chỉ tiêu duy nhất áp dụng cho các mẫu hoá thạch. Nhưng việc sử dụng hình thái đơn thuần đôi khi khó áp dụng cho những loài có hình thái biến động, thậm chí con đực và con cái của một loài có thể khác nhau đến mức dễ bị coi là hai loài khác nhau. Ngược lại, có thể nảy sinh vấn đề khi phân loại những loài rất giống nhau về bộ khung xương như cừu và dê. Ngoài ra, còn có quan điểm **loài sinh thái**. Mỗi loài sử dụng những nguồn dinh dưỡng (hay ổ **sinh thái**) riêng và thực hiện nó một cách hiệu quả trong những điều kiện nhất định của **môi trường sống**. Hai loài không thể cùng chia sẻ một ổ sinh thái trong cùng một khu vực địa lý, vì loài thành công hơn sẽ chiến thắng còn loài kia sẽ bị tiêu biến và tuyệt chủng. Ba quan điểm này bổ sung cho nhau và được thống nhất để tiến hành phân loại các loài, ngoại trừ việc áp dụng chúng cho các loài có hình thái biến động hoặc các loài có quan hệ họ hàng gần.

Các cơ chế cách ly sinh sản

Các loài khác nhau bị ngăn cản về khả năng giao phối và trao đổi vật liệu di truyền do các cơ chế cách ly. Các cơ chế này nhất thiết phải diễn ra trong quá trình hình thành loài. Các cơ chế cách ly được liệt kê ở đây dựa trên cách phân loại của Dobzhansky.

Các cơ chế cách ly trước giao phối:

Các cơ chế cách ly trước giao phối ngăn cản sự giao phối giữa các quần thể và rất hiệu quả để hình thành loài.

(i) **Cách ly tập tính** (**giới tính** hoặc **phong tục**). Các cá thể chọn các cá thể cùng loài chứ không chọn các cá thể khác loài làm bạn giao phối. Ở các loài thụ tinh ngoài, tập tính này được thể hiện bởi giao tử.

(ii) **Cách ly không gian** (**địa lý, sinh thái** hoặc **môi trường sống**). Các cá thể của những loài khác nhau không gặp gỡ nhau khi giao phối vì chúng giao phối ở những nơi hoặc môi trường khác nhau. Việc phân biệt cơ chế này dựa trên mức độ vận động của sinh vật. Các đặc điểm không gian và môi trường khác nhau có thể bao gồm châu lục, hệ sinh thái, cây cung cấp thức ăn, mức độ khô hạn, mức độ acid, sự khác biệt nhiệt độ ở các môi trường sống, v.v... Hiệu quả của cơ chế này luôn giống nhau cho dù nguyên nhân là chọn lọc theo môi trường hoặc do sự phân cắt quần thể thành những quần thể nhỏ trên các đảo, hồ, núi... do nguyên nhân địa lý, địa chất. Những côn trùng ít di chuyển trên cây này có thể bị cách ly một cách hiệu quả với quần thể trên cây khác chỉ cách nó 100m.

(iii) **Cách ly thời gian.** Các loài khác nhau giao phối hoặc nở hoa ở những thời điểm khác nhau trong ngày hoặc trong năm.

(iv) **Cách ly cơ học.** Cơ quan sinh dục ngoài hoặc các bộ phận của hoa cản trở sự giao hợp hoặc nảy mầm của hạt phấn.

(v) **Cách ly do các nhân tố thụ phấn khác nhau.** Các loài mặc dù có quan hệ gần gũi, nhưng có thể thu hút các nhân tố (ví dụ: các loài côn trùng) thụ phấn khác nhau.

(vi) **Cách ly giao tử.** Ở các loài giao phối ngoài, các giao tử có hiện tượng hấp dẫn chọn lọc. Tinh trùng của một loài có thể chỉ bơi hướng tới tế bào trứng của chính loài đó. Ở các loài thực vật, nhụy hoa có thể ngăn cản sự xâm nhập của hạt phấn mang tính không hợp. Ở các loài giao phối trong, các giao tử đực có thể không sống được trong ống dẫn sinh dục của loài khác.

Các cơ chế cách ly sau giao phối:

Các cơ chế cách ly sau giao phối làm giảm sức sống và/hay giảm khả năng hữu thụ của các con lai thu được từ phép lai giữa hai loài với nhau. Hiện tượng này sẽ góp phần ngăn cản hai loài cùng tồn tại trong một môi trường sống (cùng khu vực phân bố), trừ khi có hiện tượng di cư, bởi vì các loài hiếm có xu hướng gặp phải những loài khác khi đi tìm kiếm bạn giao phối (giao phối nhầm), và vì vậy trong phần lớn trường hợp tạo ra con lai dị hợp tử có khả năng thích nghi thấp. Các loài càng hiếm càng có nguy cơ chọn lọc cao.

(i) **Con lai không sống sót.** Hợp tử lai không có khả năng sống sót, hoặc có sức sống giảm đi. Hậu quả này là do các sản phẩm gen của hai loài bố, mẹ không tương tác phù hợp với nhau.

(ii) **Con lai bất thụ.** Một trong hai giới tính của các con lai F1 bị bất thụ hoặc giảm khả năng hữu thụ. Hiện tượng này có thể xảy ra do sự thay đổi trình tự xuất hiện từ những đột biến ngẫu nhiên. Chẳng hạn khi ADN của nấm men *Saccharomyces* phân ly khoảng 0,6% (có thể do hiện tượng cách ly diễn ra trong 300.000 năm), hệ thống sửa chữa ghép cặp nhầm bắt đầu ngăn cản sự tiếp hợp xảy ra trong giảm phân do chúng không nhận ra được đó là những NST tương đồng. Nếu sự kết cặp không diễn ra được bình thường, thì các NST sẽ không phân ly đúng về các giao tử và hiện tượng lệch bội có thể xảy ra. Sự cách ly dẫn đến hình thành loài mới có thể diễn ra nhanh hơn do sự sắp xếp mới của NST được cố định ở một trong những quần thể cách ly. Hiện tượng này có thể gây nên hiện tượng bất thụ trong các con lai dị hợp tử do các NST không phân ly đúng trong giảm phân.

(iii) **Các con lai suy thoái.** Các con lai thế hệ F2 hoặc lai phân tích (backcross) không có khả năng sống sót, bất thụ, hoặc giảm sức sống hay khả năng hữu thụ. Điều này diễn ra do các cá thể F1 với một bộ NST hoàn chỉnh từ một trong hai bố, mẹ có khả năng sống sót, nhưng sự sắp xếp lại NST sau giảm phân tạo ra các giao tử không còn mang một hệ gen nguyên vẹn nữa. Chẳng hạn, hiện tượng này có thể diễn ra do những chuyển đoạn NST nhỏ, dẫn đến việc một số gen không còn tồn tại trên các NST tương đồng.

Sự hình thành loài

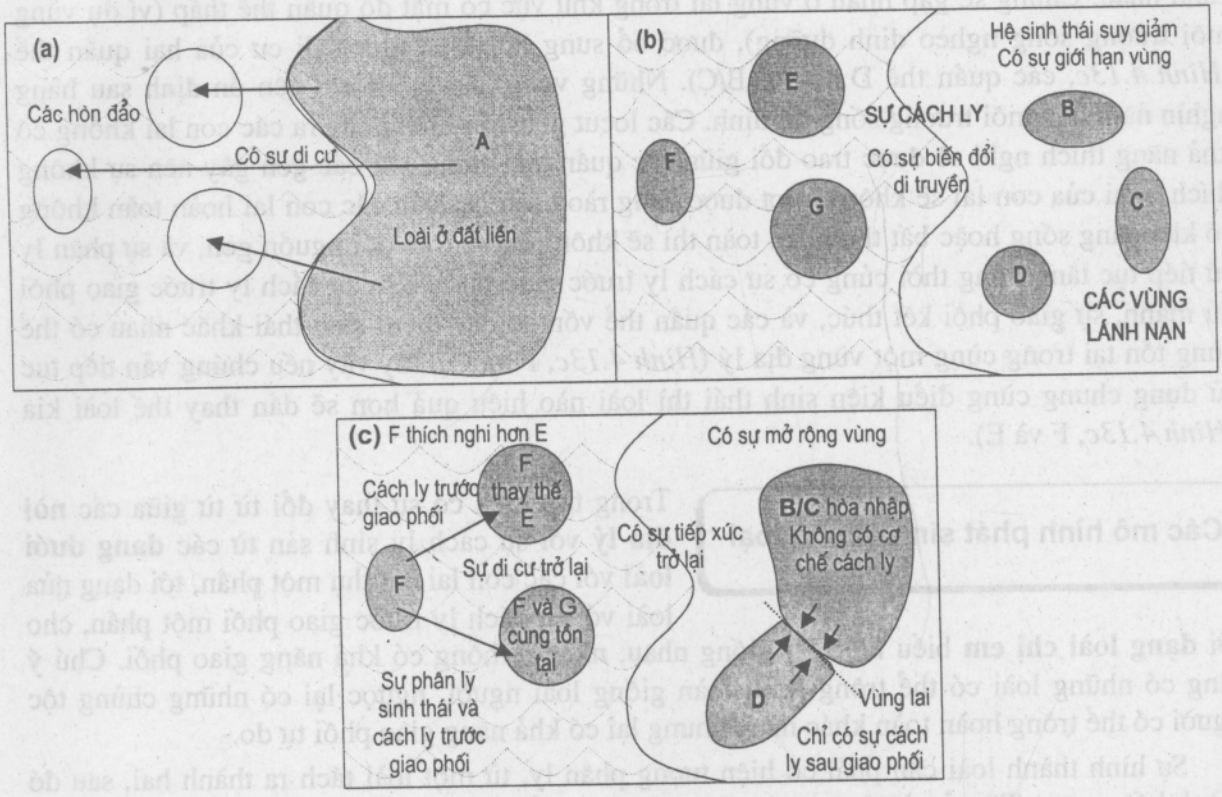
Sự hình thành loài là quá trình phân ly tạo thành hai hay nhiều loài mới xuất phát từ một loài ban đầu. Cây mô tả mối quan hệ họ hàng giữa các loài cùng nguồn gốc gọi là cây phả hệ. Các loài tách biệt chỉ có thể hình thành nếu: (i) dòng gen giữa các quần thể của một loài bị hạn chế đến mức cho phép chúng phân ly, và sau đó (ii) một thành phần của quá trình phân ly tạo ra một rào cản giao phối hoàn toàn, gọi là cơ chế cách ly giao phối. Rào cản ban đầu hình thành có thể hoặc là không gian, hoặc thời gian. Sự thay đổi khí hậu có thể làm ngăn cách các quần thể nằm trên các đỉnh núi hoặc trong các hồ cách ly, hoặc các quần thể di cư hiếm bị cách ly ở những vùng phân bố mới như trên các đảo ở Hawaii hay Galapagos (*Hình 4.13*). Các loài được hình thành do cách ly địa lý gọi là sự hình thành loài địa lý.

Sự chọn lọc gián đoạn trong các loài đa hình có thể dẫn đến sự chọn lọc hình thành hai loài nếu như những con lai trung gian của chúng có khả năng thích nghi kém hơn bố, mẹ chúng. Tuy vậy, nếu hai dạng bố, mẹ này vẫn tiếp tục giao phối thì cũng khó dự đoán bằng cách nào chúng có thể phân ly đủ để có thể hình thành nên hai loài mới.

Đôi khi quá trình chọn lọc có thể tạo ra các kiểu hình sinh thái sống trong các môi trường nằm kề nhau trong cùng một vùng địa lý. Có một số dẫn chứng chứng tỏ các quần thể trong cùng môi trường sống có thể phân ly và tiến hóa thành các loài khác nhau (ví dụ như các loài cá sống trong cùng một hồ, hay như các loài côn trùng sống dựa trên các loài thực vật khác nhau nhưng trong cùng sinh cảnh). Những loài này được gọi là **những loài đồng sinh cảnh**.

Sự thay đổi di truyền một cách ngẫu nhiên do đột biến hoặc do hiện tượng trôi dạt diễn ra trong một thời gian dài (có thể trên 300.000 năm) sẽ tạo ra hiện tượng phân ly đủ lớn trong cấu trúc ADN. Kết quả là hình thành các giao tử không có khả năng sống sót hoặc tạo ra các con lai bất thụ. Điều đó dẫn đến kết quả không thể tránh khỏi là sự hình thành loài mới nếu các quần thể bị cách ly trong một thời gian đủ lớn. Sự chọn lọc tách ly có thể gây nên những thay đổi nhanh hơn làm tăng hiệu quả phân ly. Nếu hai quần thể phát triển theo hướng thích nghi với các điều kiện môi trường khác nhau thì sự giao phối có thể tạo ra các con lai có khả

năng thích nghi kém hơn. Sau đó, do sự chọn lọc theo cơ chế cách ly trước giao phối làm ngăn cản khả năng giao phối. Điều này chỉ có thể diễn ra nếu các cơ chế cách ly sau giao phối là đủ mạnh để làm tăng sự chọn lọc và làm các cơ chế các ly trước giao phối diễn ra nhanh hơn so với dòng gen hình thành trong các con lai và tạo nên sự khác biệt cần thiết. Sự chọn lọc theo cơ chế cách ly trước giao phối cũng xuất hiện khi có sự cách ly sau giao phối xảy ra bởi một nguyên nhân nào đó (thông thường là do sự khác biệt về di truyền).



Hình 4.13. **Sự hình thành năm loài từ một loài. Một loài trên đất liền di cư ra ba hòn đảo ngoài khơi (a). Sự thay đổi điều kiện khí hậu đã làm tách biệt quần thể trên đất liền thành ba vùng lánh nạn, và phân ly thành sáu quần thể khác nhau (b). Sau đó diễn ra sự mở rộng vùng phân bố và tiếp xúc trở lại (c) mà kết quả của nó phụ thuộc vào sự cách ly giao phối và phân ly sinh thái. B và C không bị cách ly mà hòa nhập vào nhau. B/C bị cách ly sau giao phối với loài D, vì vậy B/C gặp nhau ở vùng lai. F bị cách ly trước giao phối với loài E và G. F có môi trường sống giống E nhưng thích nghi hơn, vì vậy F thay thế dần E. Sự khác biệt về sinh thái cùng giúp F và G tiến hoá song song, và chúng cùng tồn tại.**

Sự tiếp xúc trở lại

Khi các điều kiện thay đổi, các quần thể phân ly và tách ly có thể gặp lại nhau; hiện tượng này được gọi là sự tiếp xúc trở lại tại vùng tiếp xúc. Kết quả sau đó của sự tiếp xúc trở lại sẽ phụ thuộc vào mức độ mà sự cách ly sau giao phối, cách ly trước giao phối và sự phân ly sinh thái đã xuất hiện kể từ sau khi chúng phân ly. Nếu những quần thể này có khả năng giao phối tự do và các con lai có khả năng thích nghi bình thường, các quần thể đó có thể hòa nhập trở lại, tạo ra một quần tụ lai (*Hình 4.13c*, các quần thể B và C). Nếu sự giao phối tạo ra các con lai không có khả năng thích nghi, thì quá trình chọn lọc sẽ hoạt động và tăng cường cơ chế cách ly trước giao phối, bởi vì các cá thể vốn giao phối với đúng loại của chúng sẽ trở nên có ưu thế. Còn những cá thể giao phối không đúng sẽ tạo ra các con lai không có khả năng thích nghi. Những

quần thể hiếm có phổ phân bố hẹp sẽ trở nên bất lợi do khả năng giao phối "sai" với các quần thể ưu thế là rất lớn và nhiều khả năng tạo ra con lai có khả năng thích nghi kém.

Nếu sự cách ly sau giao phối yếu, và sự trao đổi vật chất di truyền (dòng gen) xảy ra thông qua các con lai, thì sự khác biệt di truyền sẽ được trung hòa, chứ không bị tăng cường tại vùng tiếp xúc, và lúc đó sự phân ly sẽ không xảy ra. Sự cách ly trước giao phối sẽ không tăng lên, nhưng hai quần thể sẽ cùng sống trong không gian kiểu cận vùng phân bố, ngay bên cạnh nhau. Chúng sẽ gặp nhau ở vùng lai trong khu vực có mật độ quần thể thấp (ví dụ vùng môi trường sống nghèo dinh dưỡng), được bổ sung bởi hiện tượng di cư của hai quần thể (*Hình 4.13c*, các quần thể D so với B/C). Những vùng lai này sẽ trở nên ổn định sau hàng nghìn năm nếu môi trường sống ổn định. Các locut gen vốn không tạo ra các con lai không có khả năng thích nghi sẽ được trao đổi giữa các quần thể, trong khi các gen gây nên sự không thích nghi của con lai sẽ không vượt được hàng rào cách ly. Nếu các con lai hoàn toàn không có khả năng sống hoặc bất thụ hoàn toàn thì sẽ không có sự trao đổi nguồn gen, và sự phân ly cứ tiếp tục tăng đồng thời củng cố sự cách ly trước giao phối. Khi sự cách ly trước giao phối đủ mạnh, sự giao phối kết thúc, và các quần thể vốn có đặc điểm sinh thái khác nhau có thể cùng tồn tại trong cùng một vùng địa lý (*Hình 4.13c*, F và G), tuy vậy nếu chúng vẫn tiếp tục sử dụng chung cùng điều kiện sinh thái thì loài nào hiệu quả hơn sẽ dần thay thế loài kia (*Hình 4.13c*, F và E).

Các mô hình phát sinh chủng loại

Trong tự nhiên có sự thay đổi từ từ giữa các **nòi địa lý** với sự cách ly sinh sản từ các **dạng dưới loài** với các con lai bất thụ một phần, tới dạng nửa loài với sự cách ly trước giao phối một phần, cho tới **dạng loài chị em** biểu hiện rất giống nhau, nhưng không có khả năng giao phối. Chú ý rằng có những loài có thể trông hoàn toàn giống loài người, ngược lại có những chủng tộc người có thể trông hoàn toàn khác nhau nhưng lại có khả năng giao phối tự do.

Sự hình thành loài cần phải có hiện tượng phân ly, từ một loài tách ra thành hai, sau đó thành bốn, v.v... Tất cả các thuật toán được sử dụng để xác định cây chủng loại phát sinh làm tăng cường một cách nhân tạo chuỗi các sự kiện phân ly, nhưng có thể gây nên một ấn tượng nhầm lẫn. Rất có thể rằng các loài không chuyên biệt với sự phân bố rộng lớn về môi trường sống và địa lý có thể tồn tại ở nhiều **vùng phân bố lánh nạn** hoàn toàn cách biệt nhau trong những điều kiện trái ngược. Ví dụ như thời kỳ băng hà đã tạo ra nhiều chỏm băng trên các vùng núi và cách ly các quần thể nằm trong những thung lũng cạnh nhau ở hai bên sườn núi. Các quần thể đó sau đó bắt đầu tách ly khỏi nhau trở thành những loài phôi thai và sẵn sàng phát tán khi điều kiện môi trường khí hậu trở nên thuận lợi. Khi một sự cố xảy ra làm tiêu diệt nhiều loài (ví dụ như sự va đập của một thiên thạch nào đó đã tuyệt diệt các loài khủng long), các thế hệ con cháu của các loài còn tồn tại có thể tách ly trong nhiều môi trường sống mới độc lập với nhau trong quá trình **phát xạ thích nghi** khi điều kiện môi trường thuận lợi hơn. Các quần thể tách ly trong các điều kiện sống mới có thể đồng thời phân ly thành những loài mới. Quá trình này tạo ra những tập hợp các loài chị em trong một thời kỳ địa chất tương đối ngắn và có thể không hoàn toàn trùng khớp với các cây chủng loại phát sinh được xây dựng theo thuật toán.

Khi các dòng dõi của một loài duy nhất thay đổi qua thời gian, nó sẽ tạo ra hàng loạt các loài niên đại trong các dấu vết hóa thạch, nhưng việc xác định các loài này lại mang tính chất tùy tiện, và thường bị ảnh hưởng bởi dấu vết còn sót lại của các hóa thạch.

IV.H ĐA BỘI THỂ

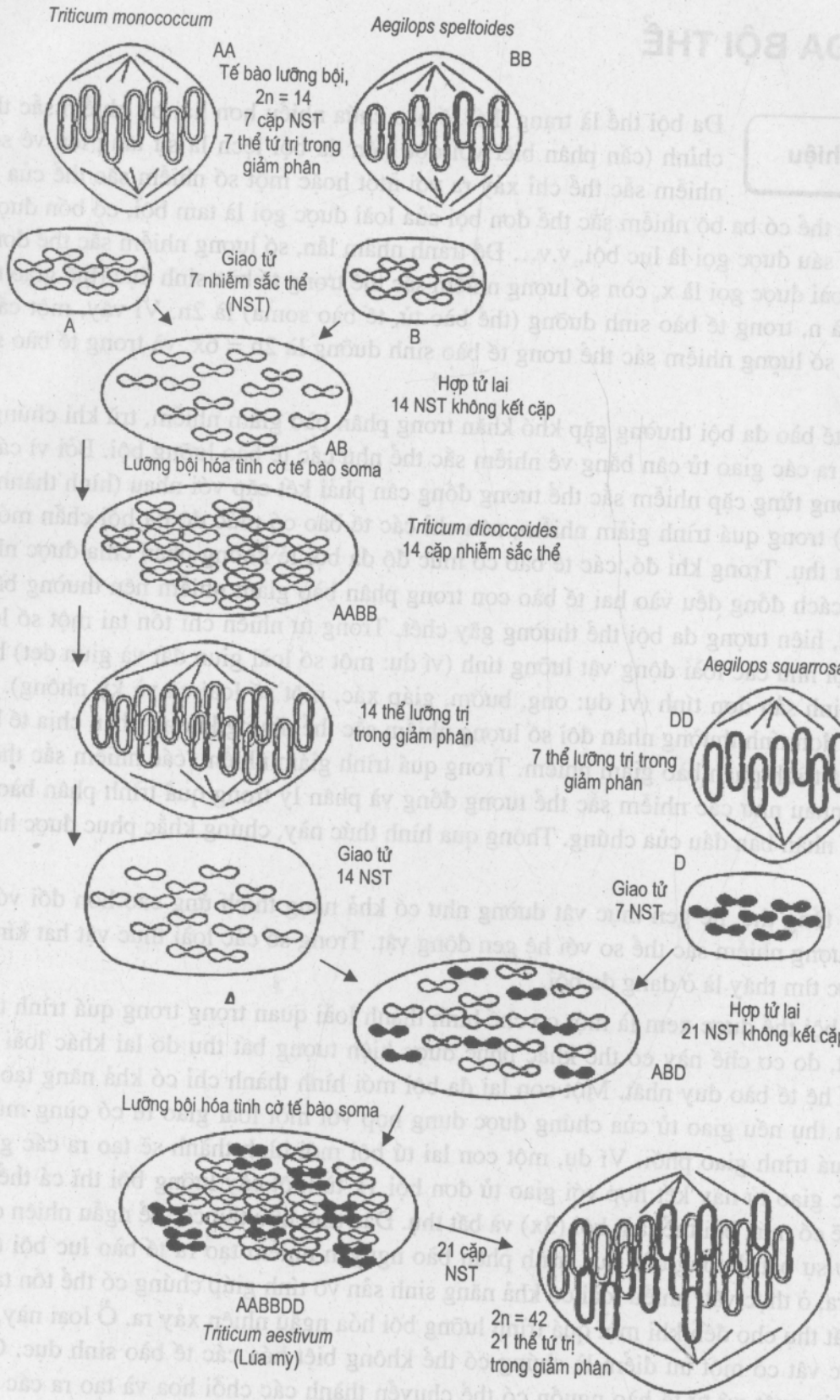
Giới thiệu

Đa bội thể là trạng thái tế bào chứa nhiều hơn hai bộ nhiễm sắc thể hoàn chỉnh (cần phân biệt với đột biến đa bội lệch là sự tăng lên về số lượng nhiễm sắc thể chỉ xảy ra với một hoặc một số nhiễm sắc thể của tế bào). Những cá thể có ba bộ nhiễm sắc thể đơn bội của loài được gọi là tam bội, có bốn được gọi là tứ bội, có sáu được gọi là lục bội, v.v... Để tránh nhầm lẫn, số lượng nhiễm sắc thể đơn bội cơ bản của loài được gọi là x, còn số lượng nhiễm sắc thể trong tế bào sinh dục (thể giao tử) được ký hiệu là n, trong tế bào sinh dưỡng (thể bào tử, tế bào soma) là 2n. Vì vậy, một cá thể lục bội sẽ có số lượng nhiễm sắc thể trong tế bào sinh dưỡng là 2n = 6x, và trong tế bào sinh dục n = 3x.

Các tế bào đa bội thường gặp khó khăn trong phân bào giảm nhiễm, trừ khi chúng có khả năng tạo ra các giao tử cân bằng về nhiễm sắc thể như các tế bào lưỡng bội. Bởi vì các nhiễm sắc tử trong từng cặp nhiễm sắc thể tương đồng cần phải kết cặp với nhau (hình thành nên thể lưỡng trị) trong quá trình giảm nhiễm, nên chỉ các tế bào có mức độ đa bội chẵn mới có khả năng hữu thụ. Trong khi đó, các tế bào có mức độ đa bội lẻ không phân chia được nhiễm sắc thể một cách đồng đều vào hai tế bào con trong phân bào giảm nhiễm nên thường bất thụ. Ở động vật, hiện tượng đa bội thể thường gây chết. Trong tự nhiên chỉ tồn tại một số loài động vật đa bội như các loài động vật lưỡng tính (ví dụ: một số loài giun đất và giun dẹt) hoặc một số loài sinh sản đơn tính (ví dụ: ong, bướm, giáp xác, một số loài cá và kỳ nhông). Các loài sinh sản đơn tính thường nhân đôi số lượng nhiễm sắc thể nhưng không phân chia tế bào ngay trước quá trình phân bào giảm nhiễm. Trong quá trình giảm nhiễm, các nhiễm sắc thể này kết cặp với nhau như các nhiễm sắc thể tương đồng và phân ly trong quá trình phân bào để phục hồi kiểu nhân ban đầu của chúng. Thông qua hình thức này, chúng khắc phục được hiện tượng bất thụ.

Xét tổng thể, hệ gen thực vật dường như có khả năng thích ứng cao hơn đối với sự thay đổi số lượng nhiễm sắc thể so với hệ gen động vật. Trong số các loài thực vật hạt kín, 47% số loài được tìm thấy là ở dạng đa bội.

Đa bội thể được xem là một cơ chế hình thành loài quan trọng trong quá trình tiến hóa ở thực vật, do cơ chế này có thể khắc phục được hiện tượng bất thụ do lai khác loài chỉ trong một thế hệ tế bào duy nhất. Một con lai đa bội mới hình thành chỉ có khả năng tạo ra thế hệ con hữu thụ nếu giao tử của chúng được dung hợp với một loại giao tử có cùng mức bội thể trong quá trình giao phối. Ví dụ, một con lai tứ bội mới hình thành sẽ tạo ra các giao tử 2x. Nếu các giao tử này kết hợp với giao tử đơn bội 1x từ một cây lưỡng bội thì cá thể con hình thành sẽ có mức bội thể tam bội (3x) và bất thụ. Đặc tính hữu thụ có thể ngẫu nhiên được phục hồi nhờ sự bất thường của quá trình phân bào nguyên nhiễm tạo ra tế bào lục bội (2n = 6x). Ngoài ra, ở thực vật, nhiều loài có khả năng sinh sản vô tính giúp chúng có thể tồn tại qua giai đoạn bất thụ cho đến khi một quá trình lưỡng bội hóa ngẫu nhiên xảy ra. Ở loại này, những cá thể thực vật có một ưu điểm là chúng có thể không biệt hóa các tế bào sinh dục. Các tế bào phân chia xuất xứ từ tế bào nguồn có thể chuyển thành các chồi hoa và tạo ra các tế bào thể giao tử có khả năng tự lưỡng bội hóa bộ nhiễm sắc thể của chúng, tạo ra hạt đa bội thể hữu thụ ở thế hệ sau.



Hình 4.14. **Quá trình tiến hóa của loài lúa mỳ lục bội *Triticum aestivum* qua quá trình lai và đa bội hóa nhắc lại**

Cũng có những quan điểm cho rằng hiện tượng đa bội hóa cũng xảy ra ở động vật có xương sống do trong hệ gen tồn tại nhiều bản sao của một gen và chúng phân bố tương đối rộng rãi. Tuy vậy, quan điểm khác lại cho rằng những trường hợp như vậy chỉ là sự sao chép của các gen ở quy mô nhỏ vốn đã xảy ra ở những sinh vật bậc thấp (chẳng hạn, hệ gen nhỏ của nấm men *Saccharomyces cerevisiae* từng có đến 50 bản sao của một gen). Những câu hỏi liên quan đến hiện tượng này có lẽ chỉ được trả lời một cách thuyết phục khi việc giải mã hệ gen hoàn chỉnh giúp khẳng định chính xác kích thước và số lượng các trình tự nguyên thủy lặp lại trong hệ gen.

Tự đa bội

Hiện tượng tự đa bội xảy ra do số lượng bộ NST trong tế bào của một cơ thể tăng lên gấp đôi (chẳng hạn, do sự phân ly NST không đồng đều trong giảm phân). Trong phần lớn trường hợp, các cá thể này cũng bất thụ do chúng có nhiều hơn hai bản sao của mỗi NST trong mỗi tế bào, dẫn đến sự kết cặp và phân ly không bình thường của các NST trong giảm nhiễm.

Dị đa bội

Các cơ thể dị đa bội có nguồn gốc con lai. Chúng được tạo ra từ sự dung hợp các tế bào sinh dục của các loài có quan hệ gần gũi. Hiện tượng này xảy ra phổ biến ở thực vật. Hai bộ nhiễm sắc thể trong con lai lưỡng bội thường khác biệt rõ rệt đến mức chúng không thể kết cặp trong kỳ đầu của quá trình giảm nhiễm. Điều đó ngăn cản việc phân ly của chúng một cách chính xác về hai cực trong lần phân bào thứ nhất của quá trình giảm nhiễm và sự phân ly chỉ thường xảy ra một cách ngẫu nhiên. Vì vậy, các giao tử có thể chứa nhiều bản sao của một số nhiễm sắc thể nào đó, trong khi lại thiếu bản sao của một hay nhiều nhiễm sắc thể khác, dẫn đến việc các giao tử không hình thành hoặc hợp tử sau khi hình thành không có khả năng sống sót. Nhưng nếu quá trình phân bào nguyên nhiễm xảy ra tình cờ cùng với việc không có sự phân chia nhân về hai tế bào con, thì số nhiễm sắc thể trong tế bào có thể được nhân lên gấp đôi tạo thành dạng dị tứ bội. Lúc này, mỗi nhiễm sắc thể lại có khả năng kết cặp với bản sao của nó, vì vậy lúc này tế bào dị tứ bội lại có hoạt động phân bào giống với tế bào lưỡng bội, nhưng giao tử của nó lúc này lại là "lưỡng bội" 2x. Hiện tượng giống như vậy được ghi nhận trong quá trình tiến hóa của loài Lúa mỳ (*Hình 4.14*). Hai loài *Triticum monococcum* (có hệ gen AA) và *Aegilops speltiodes* (hệ gen BB), mỗi loài có 7 cặp nhiễm sắc thể (x = 7) hình thành nên con lai (AB). Con lai này sau đó tự lưỡng bội để trở thành con lai chứa 14 cặp nhiễm sắc thể (2n = 28) là loài tứ bội *Triticum duococcum* (có hệ gen là AABB). Lúc này, hai bộ gen A có thể kết cặp với nhau, cũng như vậy đối với bộ gen B, dẫn đến khả năng hình thành giao tử cân bằng AB. Sự kết cặp tương đồng đặc hiệu A-A và B-B (chứ không phải sự kết cặp không tương đồng A-B) đòi hỏi hoạt động của một locut gen đặc biệt, được gọi là *Ph* (cần cho sự kết cặp của các nhiễm sắc thể tương đồng) nằm ở vai dài của nhiễm sắc thể số 5B. Hoạt động của gen này chỉ quan sát thấy ở giống Lúa mỳ trồng, mà không quan sát thấy ở các dạng tổ tiên hoang dại của chúng. Sau đó, một giao tử kiểu AB (gồm 14 nhiễm sắc thể) từ loài *T. duococcum* (2n=28) dung hợp với một giao tử đơn bội (hệ gen D) của loài *Aegilops squarrosa* (2n = 14) để tạo thành con lai tam bội (ABD). Con lai này sau đó lưỡng bội hóa thành con lai lục bội hữu thụ AABBDD (2n=42) *Triticum aestivum* là lúa mỳ được trồng hiện nay. Như vậy, loài Lúa mỳ *Triticum aestivum* chứa hai hệ gen hoàn chỉnh của ba loài tổ tiên hoang dại.

Sự kết hợp giữa các hệ gen

Sự biến nạp gen là sự truyền các gen từ một chủng, một dòng, hay từ một loài này sang một chủng, dòng hoặc loài khác. Các con lai có thể được hình thành giữa các loài có mức độ bội thể khác nhau, nhưng hậu quả của nó thường là mức bội thể tăng lên do số lượng nhiễm sắc thể phải được nhân đôi để khôi phục tính hữu thụ. Điều này dẫn đến sự kết hợp giữa các gen từ các loài có mức bội thể thấp vào loài mới có mức độ bội thể cao hơn, nhưng quá trình này thường không đảo ngược. Đến nay, không phát hiện thấy hiện tượng các gen từ các loài có mức độ bội thể cao hơn phân ly trở lại thành các loài có mức bội thể thấp hơn.

Các phức loài đa bội

Mức độ bội thể của một nhóm loài có khả năng lai với nhau thường có xu hướng tăng lên liên tục vì mỗi con lai có xu hướng lưỡng bội hóa số nhiễm sắc thể của nó để phục hồi đặc tính hữu thụ. Những nhóm loài như vậy được gọi là phức loài đa bội. Các loài thuộc họ Mộc lan Magnoliaceae là một ví dụ về hiện tượng này. Các loài lưỡng bội có số lượng bộ nhiễm sắc thể cơ bản (đơn bội) gồm từ 7 đến 10 nhiễm sắc thể (tức là gồm 7 - 10 cặp nhiễm sắc thể), trong khi những loài đã tuyệt chủng dự đoán có khoảng 5 - 6 cặp nhiễm sắc thể. Các loài tứ bội có khoảng 10 đến 15 cặp nhiễm sắc thể. Các loài này được hình thành thông qua một số con đường. Các loài có 12 cặp nhiễm sắc thể có thể là sự kết hợp của hai loài có 6 cặp NST hoặc giữa loài có 5 cặp NST với loài có 7 cặp NST. Các loài lục bội (6x) có từ 19 đến 24 cặp NST. Trong khi các loài bát bội (8x) có khoảng 24 - 25 cặp NST. Chú ý rằng 24 cặp NST có thể là kết quả của quá trình tứ bội hóa: $4 \times 6 = 24$, hay tam bội hóa: $3 \times 8 = 24$, hay $(7 + 8 = 15) + 9 = 24$. Các loài có mức bội thể 12x có thể có 38, 39, 40, 41, 43 hoặc 44 cặp NST, đặc biệt có hai loài có 57 cặp NST (có thể tương ứng với mức bội thể 18x) và 86 cặp NST (tương ứng với mức bội thể 24x).

Sự tiến hoá của các phức loài đa bội thường bắt đầu với sự có mặt của nhiều loài lưỡng bội và một số ít loài đa bội. Sau đó, các loài đa bội thường phát tán rộng và quá trình lai xảy ra phổ biến. Trải qua quá trình tiến hoá khoảng từ 500.000 đến 1.000.000 năm, những loài lưỡng bội có thể ngày càng trở nên hiếm hơn, thậm chí chỉ tồn tại trong các vùng rời rạc không có khả năng lai tạo với những loài khác. Dạng loài lưỡng bội ban đầu thậm chí có thể bị tuyệt chủng hoặc biến đổi thành một dạng khác biệt so với ban đầu. Cuối cùng, chỉ một số loài có mức bội thể cao tồn tại và duy trì mà không còn sự tồn tại của các loài có quan hệ di truyền gần gũi.

Đặc tính của các dạng đa bội thể

Tế bào của các dạng đa bội có xu hướng lớn hơn. Vì vậy, các cá thể đa bội thường có kích thước lớn và chín về mặt sinh lý chậm hơn. Phần lớn các loài hoa trồng và các cây lương thực là các dạng đa bội. Các tế bào có kích thước lớn giúp nâng cao chất lượng của các loại hoa quả và rau màu (ví dụ: râu tây, táo, khoai tây), nhờ tế bào của chúng có nhiều thành phần tế bào chất hơn so với màng và thành tế bào. Kích thước lớn cũng giúp tăng năng suất (ví dụ: lúa mỳ, bông). Hiện tượng bất thụ ở các loài tam bội được khai thác để tạo các loại quả không hạt (chẳng hạn như chuối, cam, quýt) hoặc có vị dễ ăn hơn (ví dụ: các giống dưa chuột không có vị đắng). Những con lai F1 tam bội được tạo ra bằng việc lai giữa các dạng lưỡng bội và tứ bội. Hướng ứng dụng này được áp dụng ở những loài

không thu hoạch hạt, hoặc thậm chí không muốn thu hạt (như củ cải đường, các cây ăn quả có múi), nhưng không được áp dụng ở những loài cây thu hoạch hạt (như các loài ngũ cốc chẳng hạn). Các cây chuối tam bội được nhân giống bằng phương pháp vô tính, nhưng tính bất thụ của chúng lại là trở ngại lớn trong việc chọn giống vì mục đích cải tiến năng suất và chất lượng.

Hiện tượng **đa bội hóa giúp làm ổn định hệ gen** vì sự phân ly của các kiểu gen đồng hợp tử chiếm tỷ lệ hiếm. Một con lai dị hợp tử tứ bội AAaa sẽ tạo ra các giao tử theo tỷ lệ: 1/6 AA : 4/6 Aa : 1/6 aa. Sự giao phối ngẫu nhiên cho tỷ lệ kiểu gen ở thế hệ tiếp theo sẽ là 1/36 AAAA : 8/36 AAAa : 18/36 AAaa : 8/36 Aaaa : 1/36 aaaa. Như vậy chỉ có 1/18 số cá thể là đồng hợp tử tuyệt đối, so với 1/2 trong trường hợp con lai lưỡng bội. Nếu một con lai đa bội mới hình thành tỏ ra thích nghi với điều kiện môi trường thì các thế hệ con cháu của nó cũng thường tỏ ra thích nghi với điều kiện môi trường tương tự. Tuy vậy, trong quá trình tiến hoá, các dạng đa bội thường chuyển biến chậm hơn do chúng phản ứng với các yếu tố chọn lọc chậm hơn.

IV.I QUÁ TRÌNH TIẾN HÓA

Tiến hóa do sự phân ly di truyền

Tiến hóa thực chất là sự thay đổi tần số các alen qua thời gian, với những sự thay đổi khác nhau xảy ra trong các quần thể khác nhau dẫn đến sự cách ly di truyền của các quần thể (tức là sự hình thành loài mới). Chúng ta có thể so sánh giữa các quần thể trong cùng một loài hoặc giữa các loài để nhận thấy mức độ khác nhau về mặt di truyền giữa chúng. Mức độ khác biệt về mặt di truyền giữa các quần thể và giữa các loài phản ảnh sự phân ly của chúng trong quá trình tiến hóa và là thước đo thời gian tiến hóa của chúng. Mặt khác, sự giống nhau về những đặc điểm chung giữa các quần thể và các loài có quan hệ họ hàng cho phép xác định các dạng tổ tiên nguồn gốc của chúng, là dạng mà tính trạng chung đó lần đầu tiên xuất hiện. Hai loài phân ly càng gần thì càng có nhiều đặc điểm giống nhau. Những tính trạng đó được gọi là các **tính trạng đồng dạng** (ví dụ: cấu tạo chi sau của động vật có vú). Thuật ngữ này cần phân biệt với thuật ngữ **tính trạng tương quan** (analogy) là những tính trạng mà các loài tiến hóa theo các hướng cấu trúc khác nhau nhưng có cùng chức năng (như cánh của dơi và chim chẳng hạn) thông qua quá trình **tiến hóa hội tụ**. Tuy vậy, sự phân biệt những thuật ngữ này đôi khi cũng không hoàn toàn rạch ròi. Ví dụ, tất cả các động vật có vú, bò sát và chim có kiểu cấu tạo chân trước đồng dạng được thừa hưởng từ cấu tạo chân trước của một loài lưỡng cư tổ tiên vốn bắt nguồn từ cấu tạo vây của loài cá cổ xưa hơn nữa. Tuy vậy, cánh của dơi và cánh của chim có thể coi là tính trạng đồng dạng nếu xét về phương diện chi trước, nhưng là tính trạng tương quan nếu xét đến đặc điểm tiến hóa của chúng. Trong thực tế, tổ tiên chung gần nhất của dơi là chim không có cánh.

Các quần thể

Việc so sánh giữa các quần thể trong cùng một loài cho thấy mức độ phân ly di truyền trong loài trước khi hình thành loài mới. Sự giống nhau và khác nhau có thể biểu hiện thông qua tập tính di cư, những biến đổi gây ra do áp lực chọn lọc và những biến đổi ngẫu nhiên do sự "trôi dạt" của các alen.

Điều quan trọng là phải xác định được nguyên nhân nào gây nên biến đổi gì? Quá trình chọn lọc tự nhiên ở loài người đã lựa chọn nước da sẫm cho vùng xích đạo (giả thiết là để tránh sự gây hại của ánh nắng) và nước da màu nhạt cho các vùng xa xích đạo (giả thiết là nhờ vậy ánh sáng có thể thâm nhập qua da kích hoạt hệ thống tổng hợp vitamin D và tránh được bệnh còi xương). Một số dẫn chứng chứng minh cho giả thuyết này là, mặc dù tất cả những người Mỹ bản địa bắt nguồn từ vùng Bắc Á và Alaska đều có chung một nước da sẫm màu khi sống ở vùng khô và nắng (ví dụ như Califonia, Nevada và Andes). Một ví dụ khác là những người mang gen dị hợp tử hồng cầu hình liềm được lựa chọn là tính trạng kháng ký sinh trùng sốt rét. Tình trạng sốt rét phổ biến tương quan với sự phân bố tần số alen hồng cầu hình liềm được quan sát thấy ở các nước thuộc vùng Tây Phi.

Tuy vậy, sự phân bố các alen quy định nhóm máu ABO thì khó giải thích hơn. Chẳng hạn, alen I^B (quy định nhóm máu B) có sự phân bố không đồng đều trên toàn cầu (*Hình 4.15*). Alen I^B hầu như không tồn tại trong cộng đồng những những người Mỹ bản địa cho thấy alen này hầu như không thâm nhập được từ Châu Á vào Mỹ và Úc. Hiện tượng này xảy ra có thể do sự "trôi dạt" ngẫu nhiên của alen I^B, hoặc bởi vì những quần thể người Mỹ bản địa đã được phân ly từ khi alen I^B còn rất hiếm. Sự phân bố phát tán của alen I^B trong phần còn lại của thế giới có thể do sự "trôi dạt" ngẫu nhiên hoặc do quá trình chọn lọc mà yếu tố chọn lọc là bệnh tật. Các kháng nguyên nhóm máu A và B thực chất là các nhóm saccharide trên bề mặt tế bào hồng cầu vốn cũng đã được tìm thấy ở vi khuẩn và một số loại virus. Nhiều loại bệnh khác nhau tỏ ra có ảnh hưởng lớn hơn đối với những người mang nhóm máu nhất định. Chẳng hạn, ở các quần thể người không được tiêm phòng dịch những người mang alen I^A có xu hướng dễ bị nhiễm bệnh đậu mùa hơn. Sự chọn lọc ban đầu xảy ra như vậy kèm theo sự "trôi dạt" ngẫu nhiên của các alen, cùng với hiện tượng di cư, chiến tranh đã từng xảy ra trên diện rộng trong quá trình tiến hóa của loài người có thể là nguyên nhân dẫn đến sự phân bố của các alen nhóm máu không đồng đều trên phạm vi toàn cầu. Sự di cư của những chủng tộc lớn người Mông Cổ vào vùng Đông Âu và Trung Đông trong những năm của thế kỷ XII và XIII có thể góp phần giải thích cho tần số alen I^B cao được tìm thấy ở những vùng này, nhưng lại không giải thích được cho việc xuất hiện với tần số cao của alen này ở vùng Tây Phi. Nói cách khác, đến nay các giả thiết nêu trên còn chưa được khẳng định một cách hoàn toàn chắc chắn.

Hình 4.15. **Sự phân bố của alen I^B quy định nhóm máu B (trong hệ thống nhóm máu ABO) trên toàn cầu**

Các loài địa lý

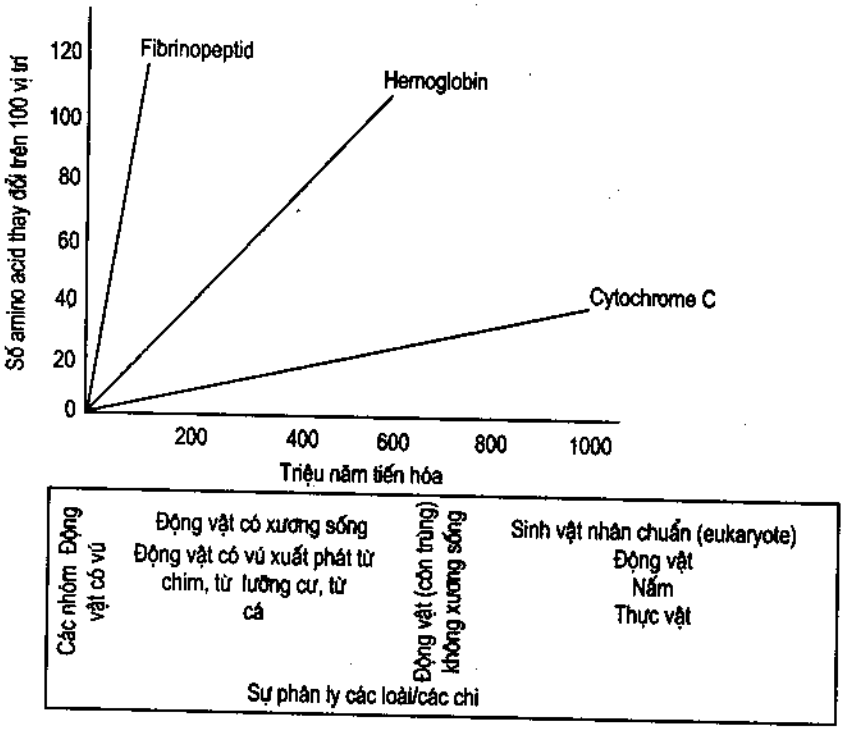
Khi một loài phân bố trên một phạm vi địa lý rộng lớn thì những quần thể xa nhau có thể phân ly về mặt di truyền đến mức chúng không còn khả năng giao phối và sinh con hữu thụ nữa, nghĩa là chúng đã phân ly thành hai loài. Những loài được hình thành như vậy được gọi là các loài địa lý. Hiện tượng này diễn ra rõ rệt nhất nếu giữa các vùng phân bố của loài có những vùng không gian cách ly không có sự cư trú của loài, vì vậy chúng còn được gọi là loài phân bố vòng tròn (ring species). Một ví dụ về loài địa lý là Mòng biển phương Bắc phân bố xung quanh Bắc Cực. Ở hai đầu tận cùng của vùng phân bố của chúng, loài Mòng biển này đã phân ly thành hai loài là *Laurus argentatus* và *L. fuscus*. Như vậy, chỉ riêng yếu tố khoảng cách địa lý cũng có thể hạn chế sự "di cư" của các dòng gen giữa các quần thể, đặc biệt việc xuất hiện vùng không phân bố của loài ở trung tâm sẽ làm cách ly các quần thể ở hai đầu, có thể dẫn đến sự hình thành loài mới. Ở một số loài chim, người ta thấy tập tính làm tổ và giao phối chỉ ở vùng nơi chúng sinh ra cũng là một nguyên nhân dẫn đến sự phân ly giữa các quần thể và khả năng hình thành loài mới, mặc dù những loài chim này thường di cư theo mùa và các quần thể không có sự ngăn cách rõ rệt về mặt địa lý.

Đồng hồ phân tử

Các đột biến tự nhiên xảy ra thường có tần số ổn định trên đơn vị thời gian và thường được coi là do sai sót của quá trình sao chép ADN. Một tỷ lệ alen đột biến hình thành có thể không gây hại (trung tính) vì vậy có thể tồn tại và tần số của chúng tăng lên trong quần thể thông qua quá trình "trôi dạt" của các alen dưới tác động của quá trình chọn lọc. Tần số các alen cũng còn phụ thuộc vào mối liên hệ giữa tính bảo thủ của sản phẩm protein (hoặc ARN) do alen mã hóa dưới tác động của chọn lọc. Khi những khác biệt về kiểu gen giữa các quần thể và các loài theo thời gian ngày càng được tích lũy nhiều hơn thì mức độ phân ly ngày càng lớn. Một số loại protein (chẳng hạn như fibrinopeptid) có khả năng biến đổi khá nhanh và có thể được dùng để phân tích các loài có quan hệ gần gũi, trong khi một số loại gen khác (như cytochrome C hay rARN) vốn có tính bảo thủ cao hơn nên được dùng để phân tích các loài có quan hệ xa nhau (*Hình 4.16*). Trên quan điểm tần số đột biến của một trình tự gen là ổn định trên đơn vị thời gian, sự thay đổi của một trình tự gen nào có thể được dùng làm **đồng hồ phân tử**.

Tất nhiên, là đồng hồ thì phải có phương pháp "đặt chuẩn", và việc này được thực hiện dựa trên các dẫn chứng hóa thạch về sự hình thành loài. Đối với hemoglobin, tần số thay đổi amino acid vào khoảng 1,2 amino acid đối với một amino acid sau cứ 1 tỷ năm (hay nói cách khác tần số thay đổi của một amino acid = $1,2 \times 10^{-9}$). Tần số này đối với fibrinopeptid thì cao hơn là $8,3 \times 10^{-9}$ (nghĩa là ít bảo thủ hơn) nhưng với histon là $0,01 \times 10^{-9}$ (có tính bảo thủ cao). Tần số thay đổi về trình tự protein dường như giống nhau ở chuột và cá voi tính trên đơn vị năm chứ không phải theo đơn vị thế hệ. Tốc độ thay đổi ở các trình tự vùng không mã hóa và những trình tự tương tự trong vùng mã hóa ở chuột dường như cao hơn ở cá voi tính theo đơn vị từng thế hệ. Điều này cho thấy dường như khi điều kiện môi trường thay đổi thì tần số đột biến cũng thay đổi theo, nhưng duy trì là một hằng số khi môi trường ổn định. Nếu giả thiết này là đúng thì yếu tố chọn lọc, chứ không phải sự "trôi dạt" của các alen trung tính, là nguyên nhân chính dẫn đến sự thay thế của phần lớn các amino acid trong quá trình tiến hóa. Tuy vậy, cũng có những quan điểm không ủng hộ về khái niệm hằng số của đồng hồ phân tử.

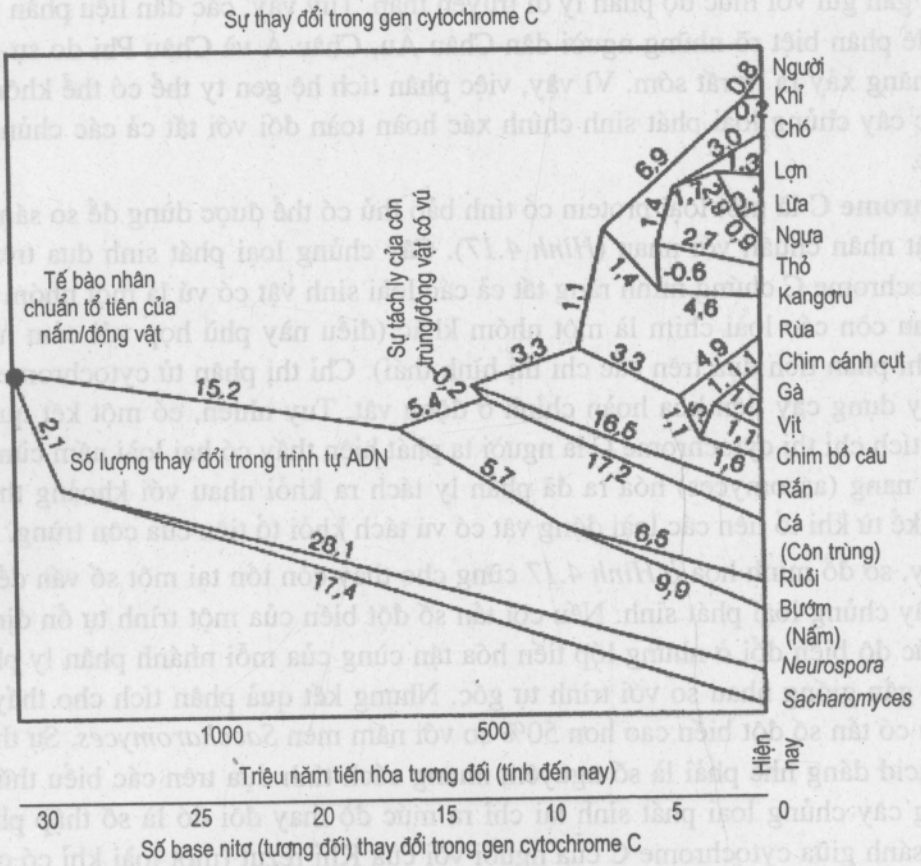
Nó có thể chạy nhanh hơn khi tốc độ tiến hóa tăng lên chẳng hạn như sự tiến hóa thích nghi ở động vật có vú sau khi các loài khủng long bị tuyệt chủng, cũng như đối với nhiều loài động vật khác vào thời kỳ cuối kỷ Phấn trắng. Đồng hồ phân tử có thể chậm hơn vào giai đoạn môi trường và quần thể ở trạng thái cân bằng. Theo đồng hồ phân tử, nhiều nhóm động vật có vú có khả năng hình thành từ sớm hơn 65 triệu năm trước, nhưng những bằng chứng hóa thạch cho thấy chúng xuất hiện trong khoảng thời gian gần hơn nhiều.



Hình 4.16. **Tốc độ thay đổi amino acid tương đối trong các protein fibrinopeptid, hemoglobin và cytochrome**

Cây chủng loại phát sinh

Cây chủng loại phát sinh được thiết lập dựa trên việc sắp xếp các loài có quan hệ di truyền gần gũi vào cùng nhóm và vẽ được đường nối phản ánh mối liên hệ giữa các loài và các nhóm loài. Khoảng cách (độ dài các đường nối) giữa các loài và các nhóm loài phản ánh mức độ biến đổi về trình tự vật chất di truyền giữa các loài và nhóm loài, qua đó có thể diễn đạt khoảng thời gian tiến hóa từ khi các loài bắt đầu phân ly. Các cây chủng loại phát sinh có thể biểu diễn dưới dạng đồ thị hình cây hoặc hình sao (*Hình 4.17*). Một vấn đề là phải tìm được điểm bắt nguồn của cây chủng loại phát sinh. Điều này có thể thực hiện được bằng việc sử dụng một loài ngoài nhóm nghiên cứu có vẻ có khoảng cách di truyền đồng đều hơn cả với tất cả các loài và nhóm loài nghiên cứu. Ví dụ, sử dụng một loài nấm hoặc thực vật làm gốc để phân tích gen cytochrome C ở động vật (*Hình 4.17*). Các trình tự có tần số đột biến cao (như ADN ty thể) được dùng để thiết lập cây chủng loại phát sinh ở những loài có quan hệ gần gũi. Những đột biến này không nên sử dụng để phân tích các loài vốn phân ly từ rất sớm trong quá trình tiến hóa vì tần số đột biến cao của chúng có thể đã tạo ra sự sắp xếp ngẫu nhiên của các trình tự. Thậm chí các đột biến phục hồi có thể đã tái tạo lại trình tự gốc ban đầu ở một số vị trí; vì vậy, kết quả so sánh trình tự trong trường hợp này có thể dẫn đến kết luận sai.



Hình 4.17. **Cây tiến hoá (cây chủng loại phát sinh) của các loài động vật dựa trên sự thay đổi nucleotid trong gen cytochrome C. Các số liệu trên sơ đồ là số nucleotid trung bình thay đổi dọc theo mỗi nhánh. Nấm được sử dụng làm gốc của cây tiến hóa**

ADN ty thể ở động vật (mADN) có tần số biến đổi nhanh hơn rất nhiều so với ADN gen nhân, có thể bởi vì kích thước hệ gen nhỏ, chịu áp lực chọn lọc thấp hơn, vì vậy quá trình sao chép ADN có xu hướng dễ gặp sai sót hơn. Đặc điểm này của hệ gen ty thể giúp việc phân tích những sự biến đổi trong thời gian ngắn dễ dàng hơn. Hơn nữa, hệ gen ty thể có đặc tính thường di truyền theo dòng mẹ và không có sự tái tổ hợp. Vì vậy, mỗi một dòng gen ty thể có sự di truyền độc lập. Việc phân tích mADN ở người cho thấy trình tự này ở tất cả mọi người trên trái đất đều bắt nguồn từ một trình tự tổ tiên ban đầu duy nhất (có thể coi đó là Êva) vốn sống cách đây khoảng 150.000 – 300.000 năm. Đáng chú ý các phương pháp toán học khẳng định các đột biến riêng biệt hoặc sự thay đổi trong các trình tự ADN không tái tổ hợp có thể truy lại đến nguồn gốc ban đầu, hay nói cách khác việc tồn tại một "Êva" có lẽ là một vấn đề không cần phải bàn cãi. Rất nhiều nhánh xuất phát từ cùng nguồn gốc này sau đó bị tuyệt chủng, có khả năng do ngẫu nhiên nhiều hơn là do quá trình chọn lọc. Tuy vậy, các dẫn chứng còn chưa thể khẳng định chắc chắn nơi cư ngụ của "Êva". Nhiều dẫn liệu cho thấy "Êva" có khả năng cư ngụ ở Châu Phi. Hai dẫn chứng ủng hộ giả thiết này là: (i) phần lớn các nhánh phân ly sau đó đều xuất hiện ở châu lục này; và (ii) phân lớn các kiểu đa hình mADN cũng được tìm thấy ở đây. Điều này cho thấy đã có các dòng gen ty thể tồn tại song song ở Châu Phi trong một thời gian dài hơn ở các châu lục khác. Phần lớn các chủng người Châu Âu có

quan hệ rất gần gũi với mức độ phân ly di truyền thấp. Tuy vậy, các dẫn liệu phân tích mADN không đủ để phân biệt rõ những người dân Châu Âu, Châu Á và Châu Phi do sự phân ly của họ có khả năng xảy ra từ rất sớm. Vì vậy, việc phân tích hệ gen ty thể có thể không đủ để có thể vẽ được cây chủng loại phát sinh chính xác hoàn toàn đối với tất cả các chủng tộc người trên trái đất.

Cytochrome C là một loại protein có tính bảo thủ có thể được dùng để so sánh tất cả các loài sinh vật nhân chuẩn với nhau (*Hình 4.17*). Cây chủng loại phát sinh dựa trên phân tích trình tự cytochrome C chứng minh rằng tất cả các loài sinh vật có vú là một nhóm có quan hệ gần với nhau còn các loài chim là một nhóm khác (điều này phù hợp với cảm nhận chung, cũng như khi phân tích dựa trên các chỉ thị hình thái). Chỉ thị phân tử cytochrome C cũng có thể giúp xây dựng cây tiến hóa hoàn chỉnh ở động vật. Tuy nhiên, có một kết quả gây ngạc nhiên phân tích chỉ thị cytochrome C là người ta phát hiện thấy có hai loài nấm cùng được xếp loại là nấm nang (ascomyces) hóa ra đã phân ly tách ra khỏi nhau với khoảng thời gian dài hơn hai lần kể từ khi tổ tiên các loài động vật có vú tách khỏi tổ tiên của côn trùng.

Tuy vậy, sơ đồ minh họa ở *Hình 4.17* cũng cho thấy còn tồn tại một số vấn đề trong việc xây dựng cây chủng loại phát sinh. Nếu coi tần số đột biến của một trình tự ổn định qua thời gian, thì mức độ biến đổi ở những lớp tiến hóa tận cùng của mỗi nhánh phân ly phải có mức độ biến đổi gần giống nhau so với trình tự gốc. Nhưng kết quả phân tích cho thấy nấm men *Neurospora* có tần số đột biến cao hơn 50% so với nấm men *Saccharomyces*. Sự thay đổi của các amino acid đáng nhẽ phải là số nguyên, nhưng cách tính dựa trên các biểu thức toán học để xây dựng cây chủng loại phát sinh lại chỉ ra mức độ thay đổi đó là số thập phân. Chẳng hạn, khi so sánh giữa cytochrome C của người với của Khỉ rezut (một loài khỉ có cùng nhánh tiến hóa với người), người ta thấy có một amino acid được thay thế. Nhưng số liệu trên cây chủng loại phát sinh cho thấy lại là 0,8 với người và 0,2 đối với Khỉ rezut. Các thuật toán đại số được xây dựng theo nguyên tắc nhị phân. Nguyên tắc đó đủ chính xác để áp dụng cho việc so sánh mADN, nhưng không chính xác hoàn toàn khi áp dụng cho hệ gen nhân vốn có khả năng tái tổ hợp cao; vì vậy, đôi khi việc sử dụng các thuật toán dẫn đến sai số phân loại các loài thành các quần thể của cùng một loài, hoặc đôi khi các quần thể trong cùng loài có hệ số khác biệt như những loài riêng rẽ.

Sự tiến hóa của loài người

Tất cả những chủng tộc người hiện nay đang tồn tại trên trái đất đều cùng một loài. Một số nghiên cứu chỉ thị ADN cho thấy mức độ khác biệt giữa các cá thể trong cùng một chủng tộc, thậm chí còn lớn hơn giữa các cá thể trung bình của các chủng tộc khác nhau. Loài người thuộc nhóm động vật có vú. Các nghiên cứu giải phẫu và sinh hóa cho thấy loài người có quan hệ gần gũi nhất với vượn người. Các dẫn liệu phân tích cho thấy vượn người phân ly từ khỉ khoảng 25 triệu năm trước. Các loài vượn mới và đười ươi phân ly khỏi vượn người tương ứng vào khoảng 10 và 8 triệu năm trước. Khoảng 4 triệu năm trước, một quần thể vượn người tổ tiên đã phân hóa tiếp thành ba nhóm, sau này tiến hóa lên thành vượn người đi hai chân, tinh tinh và khỉ gôrila. Chưa có chứng cứ xác định nhánh nào trong ba nhánh vừa nêu phân ly trước tiên trong quá trình tiến hóa.

Năm 1924, Raymond Dart phát hiện ở Nam Phi một hóa thạch vượn người có độ tuổi khoảng 2,8 triệu năm. Ông đặt tên chi này là *Australopithecus*, nghĩa là vượn người Phương

Nam. Các hóa thạch vượn người cổ xưa nhất cũng đã được tìm thấy ở Laetolil (Tanzania) và Hadar (Ethiopia). Mẫu hóa thạch vượn người cổ xưa nổi tiếng nhất là "Lucy", có đến 40% bộ xương hoàn chỉnh của một con cái được tìm thấy ở Hadar. Hóa thạch này có độ tuổi ước lượng khoảng 3,6 - 3,8 triệu năm. Hóa thạch này cho thấy, bộ xương đã có liên kết mạnh với hệ cơ, cánh tay dài và răng nanh giống với răng vượn người, nhưng đã đi bằng hai chân, lưng thẳng và ống chân giống người hiện đại. Con đực có kích thước lớn hơn con cái. Có một nhóm vượn người khác có lẽ cũng đã phân ly từ nhánh tiến hóa của loài người nhưng sau đó bị tuyệt chủng, có thể bởi kích thước lớn (*A. robustus*) hoặc bởi thân hình mảnh dẻ (có thể ở con cái). Cả hai dạng này dường như có tập tính ăn lá cây. Đặc biệt, dạng vượn người có kích thước lớn với cơ nhai phát triển mạnh và cấu tạo răng nanh và răng hàm. Điều này cho thấy những dạng vượn người này có tập tính ăn uống và điều kiện sinh thái nơi phân bố khá giống với khỉ gôrila. Các bằng chứng hóa thạch cho thấy các cá thể vượn người Phương Nam cỡ lớn cùng tồn tại với tổ tiên loài người khoảng thời gian từ 1,5 - 2 triệu năm trước.

Trong quá trình tiến hóa của loài người, việc đi bằng hai chân và giải phóng hai tay để có thể sử dụng công cụ lao động và vũ khí xảy ra rất sớm. Sau sự kiện này, kích thước hàm ngày càng giảm đi, trí thông minh và kích thước bộ não ngày càng tăng. Các kỹ năng liên quan đến việc tạo ra các công cụ lao động xuất hiện và được duy trì thông qua hình thức giáo dục nhờ sự xuất hiện ngôn ngữ và quá trình chọn lọc trí tuệ. Song song với quá trình đó, sự phối hợp giữa nhiều cá thể vượn người là cần thiết để có thể cạnh tranh được với các loài động vật ăn thịt cỡ lớn thuộc họ Mèo để trở thành những động vật ăn thịt hàng đầu, đã góp phần thúc đẩy sự phát triển của tiếng nói. Những dụng cụ lao động làm bằng đá cổ nhất tìm thấy có tuổi khoảng 2,5 triệu năm trước có liên quan đến các hóa thạch của người *Homo habilis*. Sau đó, khoảng 1,5 triệu năm trước, loài người *Homo habilis* được thay thế bởi loài người *Homo erectus* có kích thước bộ não lớn hơn và khả năng cải tiến công cụ lao động. Người *Homo erectus* di cư từ Châu Phi đến các vùng khác nhau của trái đất khoảng 700.000 năm trước. Các quần thể ở Châu Phi tiếp tục tiến hóa và hình thành loài người hiện đại, *Homo sapiens*, khoảng 100.000 đến 200.000 năm trước. Một số quần thể *Homo sapiens* sau đó di cư và thay thế dần người *Homo erectus*. Trong quá trình tiến hóa, có thể có cả những nhóm người khác là các dạng dưới loài.

Người *Nêandectan* có kích thước khá lớn với râu rậm và kích thước não lớn hơn người hiện đại. Dạng người này cũng được xếp vào *Homo sapiens*, sống cách đây khoảng từ 35.000 đến 120.000 năm, có tập tính sống theo từng gia đình và săn thú dùng bẫy. Người *Nêandectan* đã biết trông coi người ốm và hỏa thiêu người chết cùng với lễ vật cho thấy họ tin vào sự tồn tại thế giới sau khi chết. *Nêandectan* có lẽ đã là một chủng tộc người phát triển mạnh, thích nghi với các điều kiện khí hậu lạnh của thời kỳ đồ đá. Người *Nêandectan* sau đó được thay thế bằng người hiện đại xuất hiện đầu tiên ở Trung Đông khoảng 40.000 năm trước, rồi sau đó di cư tới các vùng khác nhau ở Châu Âu khoảng 35.000 năm trước. Thời điểm này cũng là thời điểm mà phần lớn các chủng tộc người hiện đại bắt đầu xuất hiện ở Châu Á và Châu Úc. Tuy vậy, đến nay còn chưa rõ vị trí chính xác của người *Nêandectan* trong sơ đồ phả hệ của vượn người. Ở Châu Phi, các loài người *H. habilis*, *H. erectus* và *H. sapiens* là các dạng tiến hóa liên tục của một loài. Các quần thể của các loài này cứ lần lượt di cư khỏi Châu Phi và thay thế những quần thể của loài trước đó ở những vùng đất thuộc vùng lục địa Á - Âu.

Phần V
CÔNG NGHỆ ADN TÁI TỔ HỢP

Phần V gồm các mục:

V.A Lai acid nucleic

Lai acid nucleic
Các mẫu dò
Thẩm tách Southern
Thẩm tách Northern
Lai insitu

V.B Tách dòng ADN

Tách dòng ADN
Các enzym giới hạn
Plasmid
Phagơ lambda (λ)
Cosmid
Nhiễm sắc thể nấm men nhân tạo (YAC)
Vector tách dòng thực vật
Các ứng dụng của tách dòng ADN

V.C PCR - phản ứng chuỗi trùng hợp

Nguyên tắc
Thành phần phản ứng PCR
Phản ứng PCR xảy ra như thế nào
Các ứng dụng của phản ứng PCR

V.D Giải mã trình tự ADN

Nguyên lý giải mã trình tự ADN
Phương pháp kết thúc chuỗi trùng hợp nhờ dideoxyribonucleotid
Giải mã trình tự ADN tự động
Phân tích trình tự ADN được giải mã

V.A LAI ACID NUCLEIC

Lai acid nucleic

Khi phân tử ADN sợi kép gặp nhiệt độ cao hoặc một số hợp chất gây biến tính, các mối liên kết hydro vốn duy trì sự liên kết giữa hai mạnh đơn bị phá vỡ. Sợi kép lúc này bị biến tính và hai mạch đơn của phân tử ADN tách nhau ra. Quá trình này được gọi là sự **biến tính của ADN**. Nếu nhiệt độ sau đó được giảm xuống, hoặc tác nhân biến tính hoá học được loại bỏ khỏi môi trường, cấu trúc ADN sợi kép được phục hồi. Quá trình thứ hai này gọi là sự **hồi tính của ADN**. Trong thực tế, các cặp phân tử acid nucleic mạch đơn nếu có phần lớn trình tự base nitơ bổ trợ với nhau đều có khả năng hình thành liên kết tạo nên phân tử acid nucleic sợi kép. Phân tử acid nucleic sợi kép này được gọi là **phân tử lai** và quá trình xảy ra được gọi là **quá trình lai**. Các phân tử lai có thể được tạo thành giữa hai mạch đơn ADN, giữa mạch đơn ADN và ARN, hoặc giữa hai mạch ARN với nhau. Đặc tính lai của các acid nucleic dựa trên nguyên tắc bổ trợ của các nucleotid được dùng trong hàng loạt các kỹ thuật phân tích nhằm phát hiện các trình tự ADN và ARN đặc thù được quan tâm nghiên cứu.

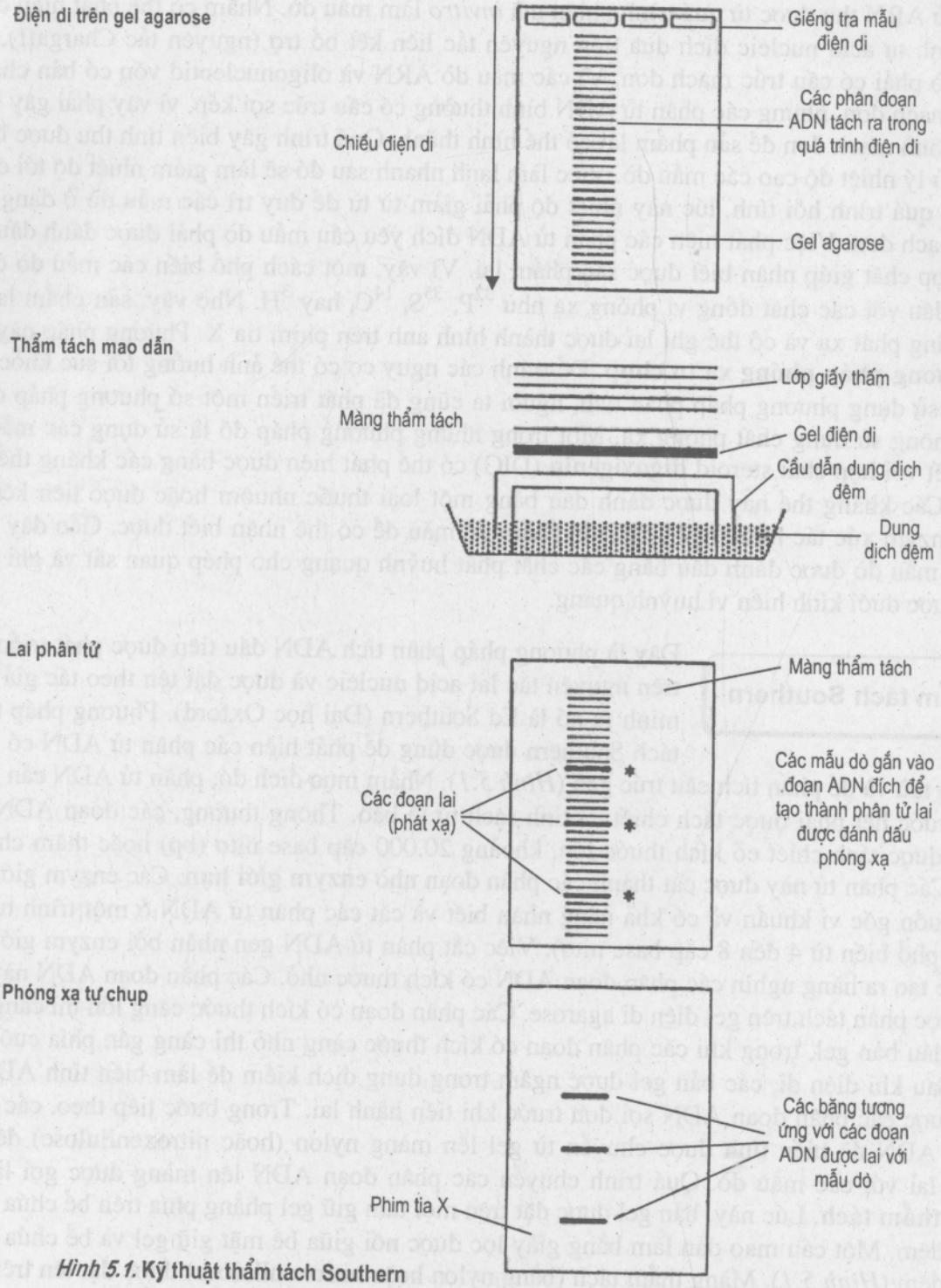
Các mẫu dò

Việc phát hiện các trình tự acid nucleic dựa trên các kỹ thuật lai phân tử cần đến sự có mặt của các mẫu dò. Các mẫu dò thực chất là một phân tử ADN hay ARN có trình tự bổ trợ với phân tử acid nucleic cần phát hiện trong mẫu nghiên cứu. Trong quá trình phân tích, các mẫu dò trước tiên phải được làm tinh sạch và không bị lẫn với các acid nucleic có trình tự khác. Các mẫu dò được sử dụng phổ biến là các trình tự ADN được tách dòng hoặc được khuếch đại nhờ phản ứng PCR. Ngoài ra, người ta cũng có thể sử dụng các đoạn oligonucleotid được tổng hợp hóa học, hoặc các phân tử ARN thu được từ quá trình phiên mã *invitro* làm mẫu dò. Nhằm có thể phát hiện được các trình tự acid nucleic đích dựa trên nguyên tắc liên kết bổ trợ (nguyên tắc Chargaff), các mẫu dò phải có cấu trúc mạch đơn. Vì các mẫu dò ARN và oligonucleotid vốn có bản chất là dạng mạch đơn, nhưng các phân tử ADN bình thường có cấu trúc sợi kép, vì vậy phải gây biến tính thành mạch đơn để sản phẩm lai có thể hình thành. Quá trình gây biến tính thu được bằng việc xử lý nhiệt độ cao các mẫu dò. Việc làm lạnh nhanh sau đó sẽ làm giảm nhiệt độ tới điểm xảy ra quá trình hồi tính, lúc này nhiệt độ phải giảm từ từ để duy trì các mẫu dò ở dạng cấu trúc mạch đơn. Việc phát hiện các phân tử ADN đích yêu cầu mẫu dò phải được đánh dấu với một hợp chất giúp nhận biết được sản phẩm lai. Vì vậy, một cách phổ biến các mẫu dò được đánh dấu với các chất đồng vị phóng xạ như ^{32}P, ^{35}S, ^{14}C, hay ^{3}H. Nhờ vậy, sản phẩm lai có khả năng phát xạ và có thể ghi lại được thành hình ảnh trên phim tia X. Phương pháp này gọi là **phương pháp phóng xạ tự chụp**. Để tránh các nguy cơ có thể ảnh hưởng tới sức khỏe của người sử dụng phương pháp phân tích, người ta cũng đã phát triển một số phương pháp đánh dấu không sử dụng chất phóng xạ. Một trong những phương pháp đó là sử dụng các mẫu dò liên kết với hợp chất steroid **digoxigenin** (DIG) có thể phát hiện được bằng các kháng thể đặc hiệu. Các kháng thể này được đánh dấu bằng một loại thuốc nhuộm hoặc được liên kết với một enzym xúc tác hình thành nên sản phẩm có màu để có thể nhận biết được. Gần đây hơn, nhiều mẫu dò được đánh dấu bằng các chất phát huỳnh quang cho phép quan sát và ghi hình ảnh dược dưới kính hiển vi huỳnh quang.

Thẩm tách Southern

Đây là phương pháp phân tích ADN đầu tiên được phát triển dựa trên nguyên tắc lai acid nucleic và được đặt tên theo tác giả phát minh ra nó là Ed Southern (Đại học Oxford). Phương pháp thẩm tách Southern được dùng để phát hiện các phân tử ADN có trình tự đặc thù và để phân tích cấu trúc gen (*Hình 5.1*). Nhằm mục đích đó, phân tử ADN cần phân tích trước hết phải được tách chiết và tinh sạch từ tế bào. Thông thường, các đoạn ADN gen nhân được tách chiết có kích thước lớn, khoảng 20.000 cặp base nitơ (bp) hoặc thậm chí lớn hơn. Các phân tử này được cắt thành các phân đoạn nhờ **enzym giới hạn**. Các enzym giới hạn có nguồn gốc vi khuẩn và có khả năng nhận biết và cắt các phân tử ADN ở một trình tự đặc hiệu (phổ biến từ 4 đến 8 cặp base nitơ). Việc cắt phân tử ADN gen nhân bởi enzym giới hạn có thể tạo ra hàng nghìn các phân đoạn ADN có kích thước nhỏ. Các phân đoạn ADN này sau đó được phân tách trên gel điện di agarose. Các phân đoạn có kích thước càng lớn thì càng gần phía đầu bản gel, trong khi các phân đoạn có kích thước càng nhỏ thì càng gần phía cuối bản gel. Sau khi điện di, các bản gel được ngâm trong dung dịch kiềm để làm biến tính ADN và thu được các phân đoạn ADN sợi đơn trước khi tiến hành lai. Trong bước tiếp theo, các phân đoạn ADN đã biến tính được chuyển từ gel lên màng nylon (hoặc nitroxenllulose) để tiến hành lai với các mẫu dò. Quá trình chuyển các phân đoạn ADN lên màng được gọi là quá trình **thẩm tách**. Lúc này, bản gel được đặt trên một tấm giữ gel phẳng phía trên bể chứa dung dịch đệm. Một cầu mao dẫn làm bằng giấy lọc được nối giữa bề mặt giữ gel và bể chứa dung dịch đệm (*Hình 5.1*). Màng thẩm tách (bằng nylon hoặc nitroxenllulose) được đặt lên trên bản

gel, rồi đặt một tập giấy thấm lên trên màng thẩm tách. Quá trình thẩm tách diễn ra trong vài giờ, khi đó dung dịch đệm từ bể chứa được thấm hút lên theo nguyên tắc mao dẫn, bắt đầu từ cầu nối chạy qua gel và qua màng lên tới tập giấy thấm. Sau vài giờ, một bản sao chính xác về vị trí, kích thước của các băng trên gel sẽ xuất hiện trên màng thẩm tách. Lúc này, màng thẩm tách được lấy ra ngoài. Các đoạn ADN trên màng sau đó được cố định bằng việc sấy ở 80°C hoặc bằng chiếu xạ tia UV.



Hình 5.1. **Kỹ thuật thẩm tách Southern**

Trong bước tiếp theo, màng được ủ với mẫu dò. Quá trình này được gọi là **quá trình lai** và được thực hiện ở điều kiện nhiệt độ và dung dịch đệm phù hợp đối với sự hình thành sản phẩm lai giữa mẫu dò và các phân đoạn ADN (lúc này đã gắn trên màng thẩm tách) theo nguyên tắc bổ trợ giữa các nucleotid. Màng lai sau đó được rửa để loại bỏ các mẫu dò thừa hoặc mẫu dò liên kết không đặc hiệu nhằm chỉ giữ lại các mẫu dò liên kết đặc hiệu với các đoạn trình tự ADN được quan tâm. Nếu như mẫu dò được đánh dấu bằng các đồng vị phóng xạ, màng lai sẽ được hiện hình trên phim tia X theo nguyên tắc phóng xạ tự chụp. Sau vài giờ, trên bản phim sẽ hiện lên các băng màu sẫm tương ứng với vị trí các đoạn ADN được lai với mẫu dò. Chiều dài các phân đoạn ADN này được tính dựa vào vị trí tương đối của chúng với thang ADN chuẩn (marker DNA) đã biết trước kích thước. Bằng cách này, người ta có thể xác định được đặc tính cấu trúc của một gen dựa trên kích thước và số lượng các phân đoạn ADN phát hiện được trong phương pháp thẩm tách Southern. Ngoài ra, phương pháp thẩm tách Southern cũng được sử dụng để phát hiện ra các thay đổi bất thường trong cấu trúc của các gen. Chẳng hạn, khoảng 20% các cá thể bị mắc bệnh máu khó đông là do sự thay đổi về cấu trúc gen tổng hợp nên protein tham gia vào quá trình đông máu, được gọi là protein yếu tố VIII. Các cá thể này có thể xác định được bằng phương pháp thẩm tách Southern. Phép phân tích có thể dùng để xác định sớm các cá thể trong gia đình có nguy cơ bị bệnh từ giai đoạn phát triển sớm của cơ thể nhằm có thể xác định các biện pháp phòng trừ bệnh thích hợp.

Thẩm tách Northern

Kỹ thuật này gần giống kỹ thuật thẩm tách Southern nhưng được sử dụng để phân tích ARN, chứ không phải là ADN. ADN trong hầu hết các tế bào của cơ thể người đều chứa các bản sao của tất cả các gen của hệ gen. Tuy vậy, ở mỗi loại tế bào cụ thể chỉ có khoảng 15% số gen hoạt động, còn 85% còn lại là không hoạt động. Những gen này không được phiên mã và vì vậy không tham gia tổng hợp protein. Ngoài ra, trong các tế bào khác nhau, sự biệt hóa hoạt động của các gen cũng khác nhau. Chẳng hạn như các gen hoạt động trong tế bào cơ khác với các gen hoạt động trong các tế bào máu. Sự biệt hóa hoạt động của các gen phản ánh thành phần protein khác nhau của các tế bào. Ở các tế bào cơ chẳng hạn, các loại protein vận động như actin và myosin vốn cần thiết cho hoạt động co giãn cơ được tổng hợp mạnh và các gen mã hóa cho chúng hoạt động rất mạnh. Ngược lại trong các tế bào máu, protein chủ yếu là hemoglobin. Lúc này, các gen tham gia tổng hợp hemoglobin hoạt động rất mạnh. Phương pháp thẩm tách Northern được dùng để xác định các gen hoạt động ở các loại tế bào khác nhau thông qua sự phiên mã của các phân tử mARN tương ứng. Các bước thực hiện của phương pháp thẩm tách Northern là tương tự như phương pháp thẩm tách Southern. Điểm khác biệt cơ bản duy nhất là vật liệu được sử dụng cho quá trình phân tích. Đối với phương pháp thẩm tách Northern, người ta sử dụng ARN được tách chiết từ một loại mô, tế bào nhất định nào đó. Hỗn hợp ARN tổng số lúc đầu sẽ chứa các phân tử mARN thu được từ quá trình phiên mã của tất cả các gen hoạt động trong loại mô, tế bào đó. Do các phân tử mARN khác nhau có kích thước khác nhau và phù hợp với kích thước của các phân tử protein tương ứng mà chúng mã hóa. Các phân tử mARN mã hóa cho các loại protein có mặt phổ biến trong tế bào sẽ chiếm ưu thế hơn so với các phân tử mARN mã hóa cho các protein ít phổ biến. Các phân tử mARN khác nhau sẽ được phân tách nhờ phương pháp điện di trên gel agarose. Mỗi một loại phân tử ARN sẽ định vị tại một vị trí nhất định trên gel điện di phụ thuộc vào kích thước của nó. Các phân tử mARN càng lớn càng nằm gần phía đầu của bản gel, trong khi các phân tử mARN càng nhỏ càng gần phía cuối bản gel. Giống như trong

phương pháp thẩm tách Southern, các gel sau đó được thẩm tách để chuyển các phân tử mARN sang màng lai và tiến hành lai với các mẫu dò đặc hiệu tương ứng với phân tử mARN được nghiên cứu. Màng lai sau đó cũng được rửa và chụp bằng phim tia X. Phim này sau khi hiện hình sẽ cho thấy các băng riêng biệt tương ứng của phân tử mARN lai với mẫu dò. Vị trí tương đối của các băng trên gel có thể dùng để ước tính chiều dài của phân tử mARN và mức độ đậm, nhạt của băng có thể dùng để định lượng phân tử mARN tương ứng có mặt trong tế bào. Trong đó, các loại phân tử mARN có hàm lượng thấp sẽ xuất hiện với các băng điện di nhạt, trong khi các loại mARN có hàm lượng cao sẽ xuất hiện bằng các băng điện di đậm. Nhờ vậy, phương pháp thẩm tách Northern không những có thể cung cấp thông tin hữu ích về các gen hoạt động trong từng loại mô, tế bào; mà còn phản ánh mức độ biểu hiện của gen và kích thước của các phân tử mARN tương ứng của chúng.

Lai *in situ*

Sự khác biệt của kỹ thuật lai *in situ* là phương pháp này được sử dụng để phát hiện các trình tự acid nucleic có mặt trong các tế bào nguyên vẹn. Ứng dụng chủ yếu của phương pháp lai *in situ* nhằm xác định mức độ biểu hiện của một loại phân tử mARN đặc trưng nào đó trong từng loại tế bào riêng biệt thuộc các mô chứa nhiều loại tế bào khác nhau. Chẳng hạn, phương pháp lai *in situ* được dùng để xác định xem liệu phân tử mARN mã hóa cho insulin có phải chỉ biểu hiện trong tế bào β của tuyến tụy hay không, v.v... Trong kỹ thuật này, các khối mô nghiên cứu được cắt thành các lát mỏng bằng thiết bị gọi là máy cắt tiêu bản hiển vi (microtom) và đặt lên trên lam kính hiển vi. Các mẫu dò được bổ sung vào tiêu bản tế bào trên lam kính và được hấp thụ vào trong tế bào chất. Ở đây mẫu dò sẽ liên kết với các phân tử mARN tương ứng để hình thành nên phức hợp lai acid nucleic. Phức hợp lai này có thể dễ dàng phát hiện nhờ chất đánh dấu của mẫu dò và biểu hiện thành các vùng nhuộm màu trong các tế bào khi được quan sát dưới kính hiển vi. Các mẫu dò có thể được đánh dấu theo một số cách khác nhau. Chẳng hạn như, người ta có thể sử dụng các chất đồng vị phóng xạ, các chất phát huỳnh quang, các enzym xúc tác phản ứng hình thành các sản phẩm có màu, hay sử dụng steroid digoxingenin.

Một phương pháp cải tiến của kỹ thuật này gọi là **phương pháp lai huỳnh quang *in situ*** (FISH = fluorescent *in situ* hybridization) sử dụng việc lai mẫu dò với các nhiễm sắc thể trong tế bào nhằm xác định vị trí của một gen nào đó trên một nhiễm sắc thể. Trong phương pháp này, các tế bào đang phân chia ở kỳ giữa chứa các nhiễm sắc thể ở giai đoạn có ngắn cực đại với hình dạng xác định được sử dụng để tiến hành lai. Bằng việc tiến hành lai với các mẫu dò, người ta có thể xác định được vị trí của locut tương ứng trên nhiễm sắc thể khi quan sát tế bào dưới kính hiển vi huỳnh quang.

V.B TÁCH DÒNG ADN

Tách dòng ADN

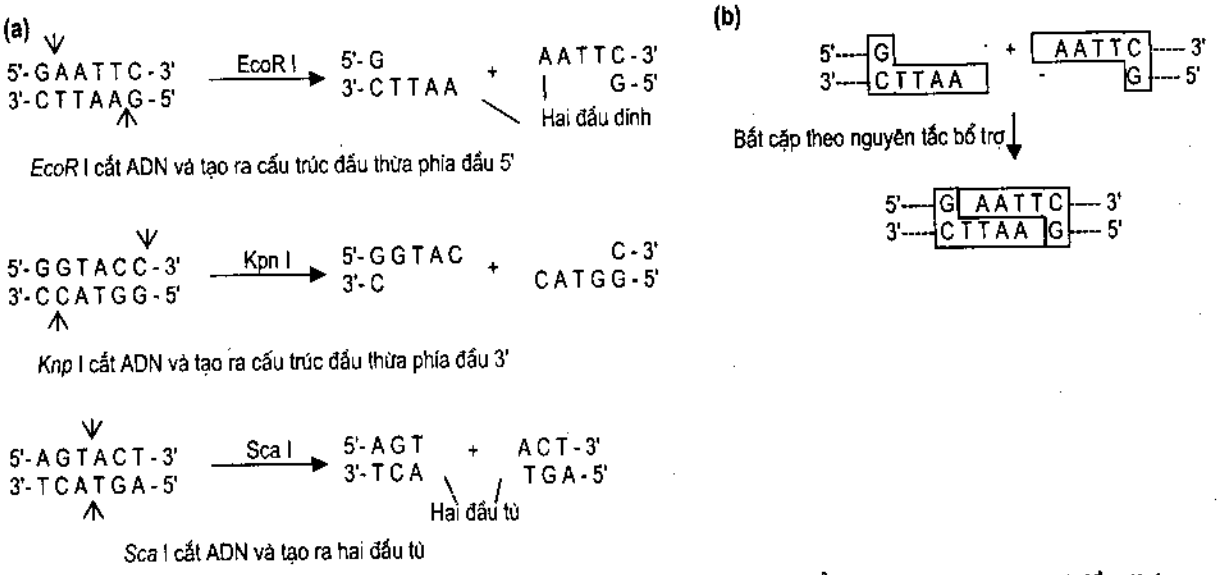
Hệ gen người ước tính có khoảng 30.000 - 40.000 gen. Trong số ADN được phân lập từ các tế bào người, các trình tự gen đơn lẻ chỉ chiếm một tỷ lệ nhỏ. Việc tách dòng ADN là một kỹ thuật hiệu quả cho phép khuếch đại một số lượng lớn bản sao của một trình tự ADN xác định nào đó để có thể sử dụng cho việc phân tích hoặc cho các mục đích thao tác di truyền sau đó. Một trong

những ứng dụng quan trọng của kỹ thuật tách dòng ADN là việc phân lập các gen mới để tiến hành nghiên cứu và phân tích chúng được chi tiết hơn.

Tất cả các thí nghiệm tách dòng ADN được dựa trên việc tạo ra các phân tử **ADN tái tổ hợp**. Quá trình này bao gồm việc tạo liên kết giữa các phân tử ADN có nguồn gốc khác nhau với nhau. Phân tử ADN được tách dòng (thường là một đoạn ADN được quan tâm) sẽ được gắn vào một phân tử ADN khác thường có cấu trúc vòng được gọi là **vector**. Vector tái tổ hợp sau đó thường được đưa vào tế bào vật chủ, phổ biến nhất là vi khuẩn *E. coli*, để chúng tự sao chép thành nhiều bản sao. Khi tế bào chủ phân chia, các bản sao của vector tái tổ hợp sẽ được chuyển vào các tế bào con. Bằng cách đó, người ta có thể thu và tinh sạch được một số lượng lớn vector từ các tế bào vật chủ được nuôi cấy và sau đó sử dụng cho việc phân tích đoạn gen tái tổ hợp quan tâm.

Các enzym giới hạn

Khả năng gắn kết các phân tử ADN khác nhau với nhau để tiến hành tách dòng phụ thuộc vào việc sử dụng các enzym có nguồn gốc vi khuẩn được gọi là các enzym giới hạn (restriction endonucleases). Các enzym này cắt các phân tử ADN tại các trình tự nhất định, thông thường được xác định bởi từ 4 đến 8 base nitơ. Các trình tự được dùng để nhận biết này có đặc điểm là trình tự đọc xuôi và ngược như nhau, còn được gọi là **trình tự giới hạn**. Nghĩa là trình tự này nếu đọc theo chiều 5'→3' sẽ giống hệt nhau ở cả hai chuỗi. Mỗi enzym giới hạn có một trình tự giới hạn xác định. Chẳng hạn enzym ***Eco*RI**, có nguồn gốc từ vi khuẩn *E. coli*, cắt bất cứ mạch ADN nào mang trình tự GAATTC. Các enzym giới hạn khác nhau nhận biết các trình tự khác nhau. Quá trình cắt của các enzym giới hạn thường theo kiểu chữ chi (nhóm enzym giới hạn loại II), vì vậy hai mạch của chuỗi xoắn kép thường được cắt rời ở một số base nitơ. Quá trình này tạo ra các đầu nhô ra ở mỗi mạch đơn gọi là **các đầu dính** ở hai đầu của phân tử ADN. Một số enzym giới hạn cắt ADN và tạo ra đầu dính ở phía đầu 5', trong khi một số enzym giới hạn khác tạo ra đầu dính ở phía đầu 3'. Một số enzym giới hạn cắt hai mạch của chuỗi xoắn kép ở cùng một điểm tạo ra cấu trúc gọi là **đầu tù** hay **đầu bằng** (*Hình 5.2a*). Hai phân tử ADN được cắt bởi cùng một enzym giới hạn



Hình 5.2. **(a) Hoạt động của các enzym giới hạn cắt ADN. (b) Sự kết cặp bổ trợ của các base từ hai đầu dính**

sẽ có khả năng gắn kết với nhau do sự bổ trợ giữa các phần đầu dính của chúng (*Hình 5.2b*). Mặc dù các phân tử này có thể gắn kết với nhau bằng các đầu dính, nhưng chúng không có khả năng tự tạo liên kết cộng hóa trị với nhau, mà cần một enzym gắn gọi là **ADN ligase** có mặt trong mọi tế bào. Enzym này là cần thiết để xúc tác sự hình thành liên kết phosphodiester giữa hai phân tử ADN, nhờ vậy tạo nên phân tử ADN tái tổ hợp mới bền vững. Một số hệ thống đã được dùng để tách dòng các phân tử ADN dựa trên việc sử dụng các loại vector khác nhau. Mỗi hệ thống thường có các đặc điểm và khả năng ứng dụng riêng biệt.

Plasmid

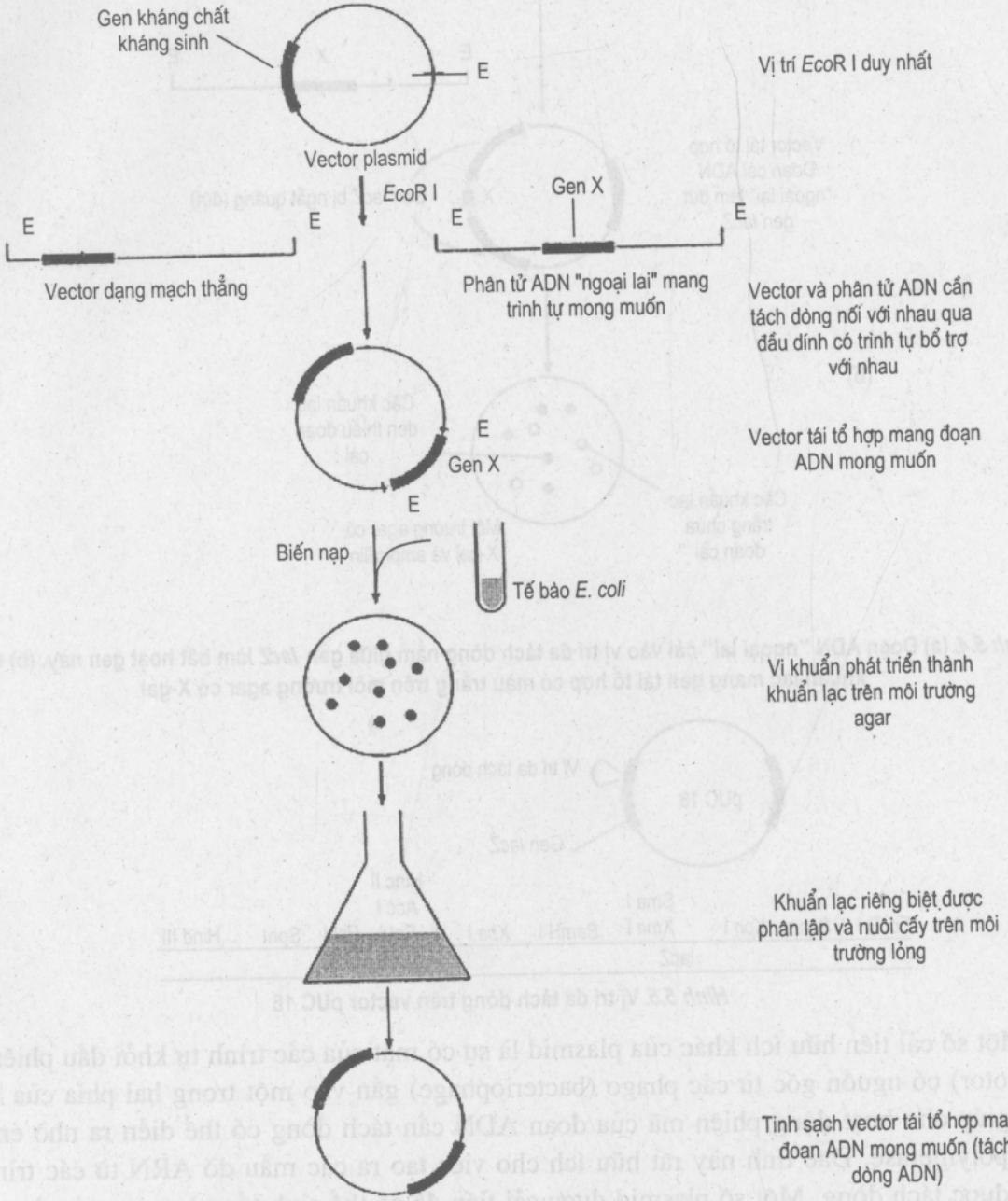
Đây là các phân tử ADN nhỏ mạch vòng có mặt phổ biến ở vi khuẩn. Chúng xuất hiện bên ngoài các nhiễm sắc thể chính của vi khuẩn và có khả năng tự sao chép. Plasmid là loại vector tách dòng được sử dụng phổ biến nhất. Quá trình tách dòng gồm có một số bước (*Hình 5.3*). Trước tiên, các plasmid được cắt bằng một enzym giới hạn tại một vị trí xác định và chuyển nó từ dạng phân tử mạch vòng thành một phân tử mạch thẳng có các đầu dính. Phân tử ADN được mong muốn tách dòng cũng sẽ được cắt bằng enzym giới hạn cùng loại để tạo ra cùng một kiểu đầu dính. Khi tiến hành trộn plasmid và phân tử ADN cần tách dòng với nhau sau khi đã được xử lý cắt bằng cùng enzym giới hạn, chúng sẽ gắn kết với nhau thông qua các đầu dính tương ứng của chúng và tạo nên các phân tử plasmid vòng tái tổ hợp mới. Enzym ADN ligase được bổ sung để tạo ra liên kết cộng hóa trị giữa hai phân tử ADN. Phân tử plasmid tái tổ hợp tiếp tục được chuyển vào tế bào vi khuẩn chủ (thường là *E. coli*) thông qua quá trình gọi là **chuyển gen** hay **biến nạp gen**. Vi khuẩn được chuyển gen sau đó được trải đều trên bề mặt môi trường thạch và khuẩn lạc chứa các tế bào mang plasmid tái tổ hợp sẽ phát triển. Các khuẩn lạc riêng rẽ sẽ được phân tách và nuôi cấy trong môi trường lỏng. Bằng cách đó, người ta có thể thu nhận và tinh sạch một số lượng lớn plasmid từ tế bào nuôi cấy và phân tử ADN được tách dòng sẽ được sử dụng cho các thí nghiệm phân tích tiếp theo.

Một số đặc tính của các plasmid là phù hợp với vai trò các vector tách dòng. Kích thước nhỏ của chúng (thường khoảng 3kbp) giúp cho việc phân lập và tách chiết từ các tế bào vi khuẩn nuôi cấy được dễ dàng, nhờ vậy có thể tái thu hồi các phân tử ADN được tách dòng. Hơn nữa, các plasmid thường mang các gen mã hóa protein kháng thuốc kháng sinh của vi khuẩn như ampicillin hay tetracyclin. Bằng việc nuôi cấy các khuẩn lạc trên môi trường agar chứa chất kháng sinh, người ta có thể nhận biết được các dòng vi khuẩn đã mang các plasmid tái tổ hợp thành công bởi vì chỉ các dòng vi khuẩn này có khả năng kháng chất kháng sinh và sinh trưởng được trên môi trường chọn lọc.

Các vector tách dòng plasmid đầu tiên dựa trên các plasmid xuất hiện trong tự nhiên. Các vector này sau đó dần dần được thay thế bằng các vector cải tiến với các trình tự ADN đã được biến đổi để có thêm một số đặc tính mới phù hợp cho các mục đích tách dòng ADN.

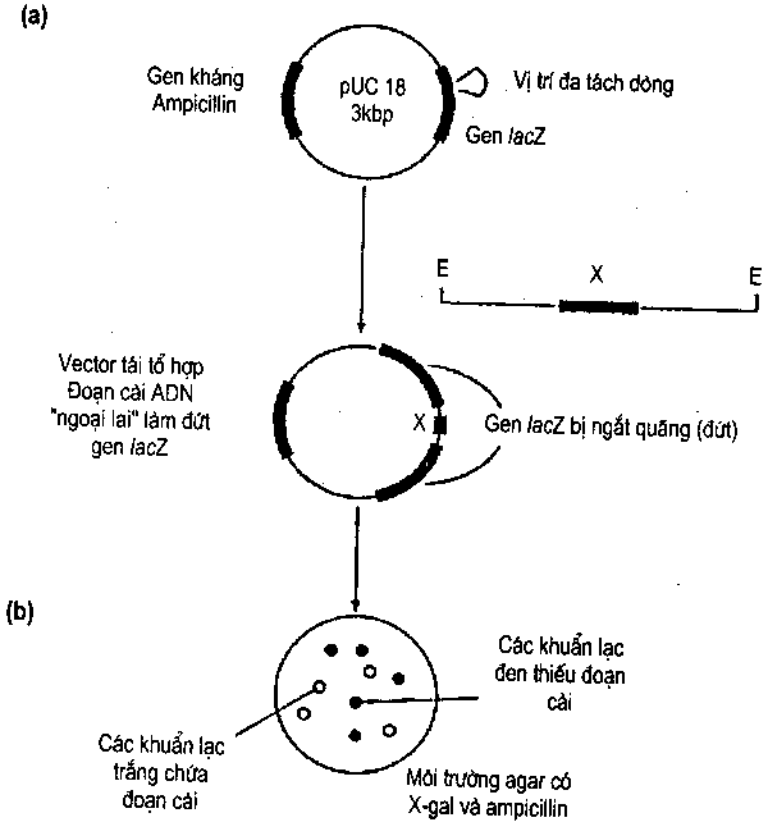
Một trong những vector plasmid thế hệ đầu tiên là pBR322. Plasmid này chứa hai gen kháng kháng sinh **ampicillin** và **tetracyclin**. Trong quá trình tách dòng, phân tử ADN khác nguồn gốc được gắn vào vị trí của gen kháng tetracyclin và vì vậy bất hoạt gen này. Vi khuẩn chuyển gen mang plasmid tái tổ hợp sau đó có thể được xác định bằng tính kháng với ampicillin nhưng mẫn cảm với tetracyclin. Các vi khuẩn mang plasmid nhưng không có đoạn ADN khác nguồn gốc có thể xác định được bằng khả năng kháng của chúng với cả hai chất kháng sinh.

Tiếp theo việc phát triển plasmid pBR322 là một loạt các vector thuộc họ pUC cho phép xác định được các khuẩn lạc mang các plasmid tái tổ hợp bằng phương pháp gọi là chọn lọc **xanh/trắng** (*Hình 5.4*). Phương pháp này dựa trên sự có mặt của gen gọi là *lac Z* mã hóa cho enzym β-galactosidase xúc tác phản ứng phân giải cơ chất tổng hợp là **X-gal** (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) không màu thành sản phẩm có màu xanh. Khi các khuẩn lạc được nuôi cấy trên môi trường agar chứa X-gal chúng sẽ có màu xanh. Tuy vậy, khi đoạn ADN "ngoại lai" gắn vào plasmid pUC, gen *lac Z* bị đứt quãng và vì vậy không sinh tổng hợp được β-galactosidase. Hậu quả là các khuẩn lạc mang plasmid tái tổ hợp sẽ có màu trắng khi nuôi cấy trên môi trường có X-gal và dễ dàng phân biệt được với các khuẩn lạc mang plasmid nhưng không có gen ADN cần tách dòng nhờ màu xanh của loại khuẩn lạc này.

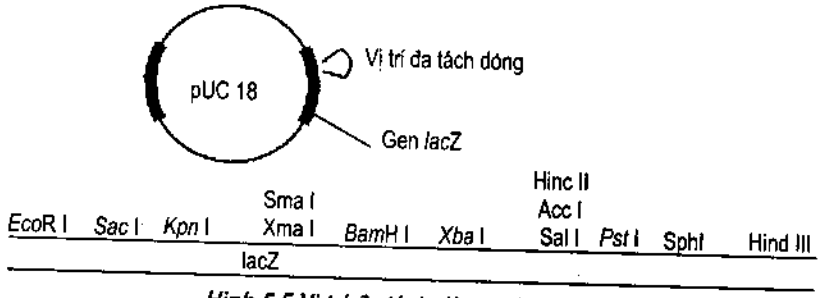


Hình 5.3. **Tách dòng ADN bằng vector plasmid**

Một đặc tính hữu ích khác của các vector pUC là trình tự của phần gen *lac Z* đã được cải tiến để hình thành một loạt các vị trí cắt của các enzym giới hạn khác nhau. Vị trí này gọi là **vị trí đa tách dòng** (MCS = *m*ultiple *c*loning *s*ite). Mục đích của việc này là tạo ra tính linh động của các vector có thể sử dụng để tách dòng các đoạn ADN khác nhau có các trình tự giới hạn khác nhau.



Hình 5.4. **(a) Đoạn ADN "ngoại lai" cài vào vị trí đa tách dòng nằm giữa gen *lacZ* làm bất hoạt gen này. (b) Các khuẩn lạc mang gen tái tổ hợp có màu trắng trên môi trường agar có X-gal**

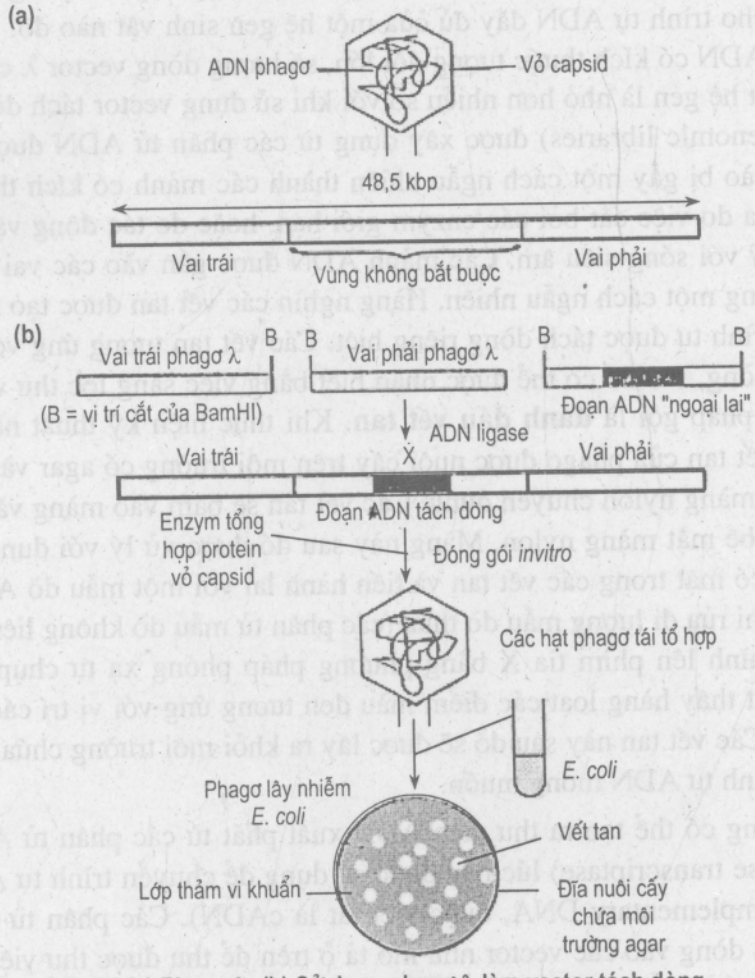


Hình 5.5. **Vị trí đa tách dòng trên vector pUC 18**

Một số cải tiến hữu ích khác của plasmid là sự có mặt của các trình tự khởi đầu phiên mã (promotor) có nguồn gốc từ các phagơ (bacteriophage) gắn vào một trong hai phía của vị trí MCS giúp cho hoạt động phiên mã của đoạn ADN cần tách dòng có thể diễn ra nhờ enzym ARN polymerase. Đặc tính này rất hữu ích cho việc tạo ra các mẫu dò ARN từ các trình tự ADN được tách dòng. Một số plasmid được cải tiến để có thể sinh tổng hợp protein được mã hóa bởi đoạn ADN được tách dòng. Các vector này được gọi là các **vector biểu hiện.**

Phagơ lambda (λ)

Thực khuẩn thể (bacteriophage/phagơ) là các virus lây nhiễm ở vi khuẩn. Chúng có hệ gen acid nucleic được bao bọc bên trong lớp vỏ protein bảo vệ gọi là protein vỏ virus (capsid). Thực khuẩn thể λ lây nhiễm ở *E. coli* có hệ gen là ADN sợi kép mạch thẳng (xem mục II.H). Phagơ tự nó gắn lên bề mặt vi khuẩn và "tiêm" ADN của nó vào trong tế bào vi khuẩn. Bên trong tế bào, đoạn ADN của phagơ λ sẽ tự sao chép nhiều lần và đồng thời sinh tổng hợp protein vỏ virus. ADN virus sau đó được đóng gói trong protein vỏ của nó và các hạt virus mới được tạo thành và giải phóng gây nên sự phá vỡ và tiêu diệt tế bào vi khuẩn bị lây nhiễm.



Hình 5.6. **(a) Phagơ λ. (b) Sử dụng phagơ λ làm vector tách dòng**

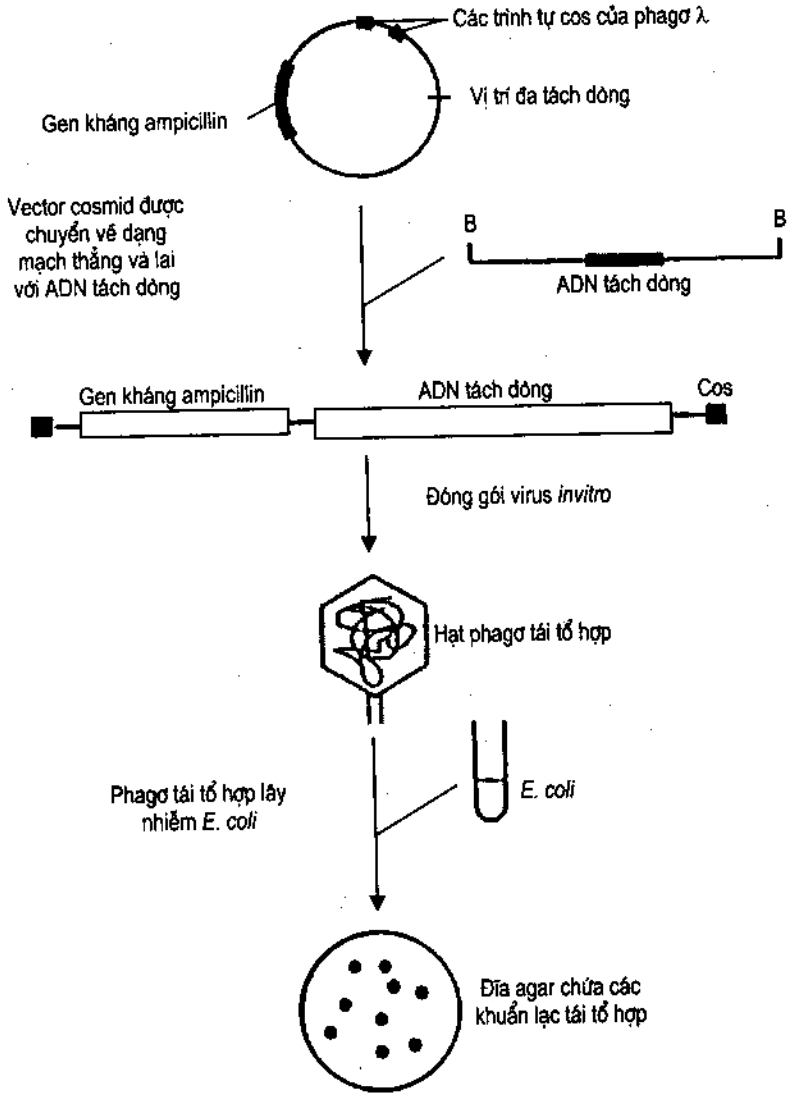
Phagơ λ đã được cải tiến để sử dụng làm vector tách dòng. Phần trung tâm của ADN phagơ λ không giữ vai trò quan trọng trong quá trình lây nhiễm được cắt bỏ để lại hai đoạn ở phần đầu 3' và 5' được gọi là **các vai** (vai trái và vai phải; xem *Hình 5.6a*). Vùng ADN của phagơ sau khi cắt bỏ được thay thế bằng phân tử ADN khác nguồn gốc cần tách dòng để hình thành nên phân tử ADN phagơ tái tổ hợp (*Hình 5.6b*). Phân tử ADN phagơ tái tổ hợp sau đó được lắp ráp vào vỏ của phagơ trong điều kiện *invitro* được gọi là quá trình **đóng gói *invitro*** bằng việc trộn phân tử ADN phagơ tái tổ hợp với dịch chiết đóng gói chứa các protein vỏ và enzym của virus. Các hạt phagơ tái tổ hợp được tạo ra có hiệu quả lây nhiễm *E. coli* cao. Các tế bào lây nhiễm được trải rộng trên bề mặt môi trường nuôi cấy có agar và chúng sinh trưởng

tạo ra **một lớp tế bào** vi khuẩn chứa các vùng trong suốt có kích cỡ bằng đầu kim. Những vùng này tương ứng với sự phân hủy vi khuẩn do sự lây nhiễm của virus và được gọi là các vết tan. Các **vết tan** riêng biệt sau đó được phân lập và dùng để tạo ra một số lượng lớn các phân tử ADN tách dòng bằng việc lây nhiễm vào các tế bào *E. coli* mới.

Ưu điểm chính của vector tách dòng phagơ λ là ở chỗ kích thước các đoạn ADN có thể tách dòng lớn hơn rất nhiều so với các vector plasmid. Các vector λ có thể "chứa" các đoạn ADN có kích thước tới 25kb so với kích thước ~10kb của các plasmid. Khả năng có thể tách dòng các đoạn ADN có kích thước lớn là lý do người ta sử dụng vector λ với mục đích chính là để xây dựng các **thư viện ADN**. Một thư viện là một tập hợp các phagơ tái tổ hợp chứa các dòng đại diện cho trình tự ADN đầy đủ của một hệ gen sinh vật nào đó. Nhờ phagơ λ có thể chứa các đoạn ADN có kích thước tương đối lớn, số lượng dòng vector λ cần và đủ để đại diện cho toàn bộ một hệ gen là nhỏ hơn nhiều so với khi sử dụng vector tách dòng là plasmid. **Thư viện hệ gen** (genomic libraries) được xây dựng từ các phân tử ADN được tách chiết và tinh sạch từ các tế bào bị gẫy một cách ngẫu nhiên thành các mảnh có kích thước khoảng 20kbp, hoặc được tạo ra do việc cắt bởi các enzym giới hạn, hoặc do tác động vật lý trong quá trình pipét hoặc xử lý với sóng siêu âm. Các mảnh ADN được gắn vào các vai của vector phagơ λ và được tách dòng một cách ngẫu nhiên. Hàng nghìn các vết tan được tạo ra, trong đó mỗi vết tan mang một trình tự được tách dòng riêng biệt. Các vết tan tương ứng với một trình tự được quan tâm tách dòng, sau đó có thể được nhận biết bằng việc sàng lọc thư viện hệ gen với mẫu dò nhờ phương pháp gọi là **đánh dấu vết tan**. Khi thực hiện kỹ thuật này, người ta lấy đĩa petri chứa các vết tan của phagơ được nuôi cấy trên môi trường có agar và đặt lên bề mặt môi trường một tấm màng nylon chuyên dụng. Các vết tan sẽ bám vào màng và tạo ra một bản sao các vết tan trên bề mặt màng nylon. Màng này sau đó được xử lý với dung dịch kiềm để gây biến tính ADN có mặt trong các vết tan và tiến hành lai với một mẫu dò ADN được đánh dấu phóng xạ. Sau khi rửa đi lượng mẫu dò thừa (các phân tử mẫu dò không liên kết), màng lai thu được được ghi hình lên phim tia X bằng phương pháp phóng xạ tự chụp. Trên phim tia X, người ta quan sát thấy hàng loạt các điểm màu đen tương ứng với vị trí các vết tan chứa trình tự mong muốn. Các vết tan này sau đó sẽ được lấy ra khỏi môi trường chứa agar và sử dụng để phân lập đoạn trình tự ADN mong muốn.

Người ta cũng có thể tạo ra thư viện ADN xuất phát từ các phân tử ARN. Enzym phiên mã ngược (reverse transcriptase) lúc này được sử dụng để chuyển trình tự ARN thành trình tự ADN bổ trợ (complementary DNA, được viết tắt là cADN). Các phân tử cADN được tạo ra sau đó được tách dòng vào các vector như mô tả ở trên để thu được thư viện hệ gen. Thư viện ADN được tạo ra bằng phương pháp này được gọi là **thư viện cADN** và chứa các dòng đại diện cho các gen đang hoạt động trong tế bào được sử dụng để tách chiết ARN. Vì vậy, một thư viện cADN thu được từ tế bào máu sẽ có nhiều dòng gen không giống thư viện cADN thu được từ tế bào phổi hay tế bào thận do các gen ở trạng thái hoạt động trong mỗi loại tế bào này là không giống nhau, dù chúng có nguồn gốc từ cùng một cơ thể. Các thư viện cADN có một ưu điểm là chúng không mang các trình tự intron. Nhờ vậy, việc phân lập và nghiên cứu đặc tính các gen được tách dòng là đơn giản hơn. Thông thường, một trình tự mã hóa một gen hoàn chỉnh chỉ xuất hiện trong thư viện cADN bằng một dòng duy nhất. **Thư viện biểu hiện gen** (expression library) là một loại thư viện cADN trong đó các trình tự được tách dòng được dịch mã thành các phân tử protein tương ứng nhờ các tế bào vi khuẩn chủ. Nhờ vậy, người ta

có thể sàng lọc thư viện bằng việc sử dụng các kháng thể đặc hiệu đối với phân tử protein được mã hóa bởi trình tự tách dòng.



Hình 5.7. **Tách dòng ADN bằng vector cosmid**

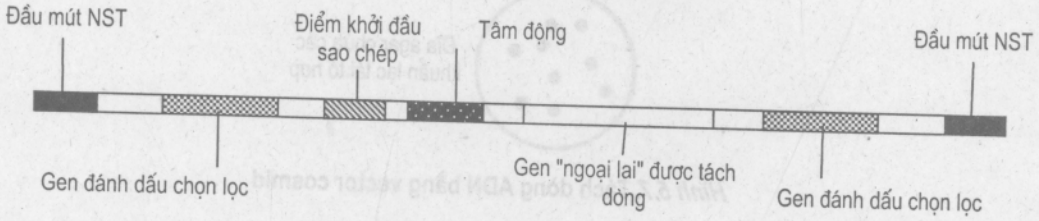
Cosmid

Loại vector này mang các đặc tính phối hợp của vector plasmid và phagơ λ. Cosmid có đầy đủ các đặc tính của vector plasmid, bao gồm một điểm đa tách dòng (MCS) và các gen kháng chất kháng sinh, nhưng đồng thời nó chứa các trình tự tìm thấy trong vector phagơ λ là các đoạn **trình tự cos**. Các trình tự cos có ở cả hai đầu của phân tử ADN của λ và có vai trò trong việc gắn phân tử của nó vào phần vỏ của phagơ. Việc có mặt trình tự cos trong cosmid giúp cho nó có thể đóng gói tự nhiên vào phần vỏ của phagơ. Việc tách dòng nhờ cosmid sẽ giúp phối hợp được các đặc tính giống như khi sử dụng cả hai vector tách dòng là λ và plasmid (*Hình 5.7*). Phân tử ADN cosmid được cắt nhờ việc sử dụng một enzym giới hạn và sau đó gắn vào phân tử ADN ngoại lai. Phân tử cosmid tái tổ hợp sau đó được đóng gói vào phần vỏ của phagơ λ, rồi được tiến hành lây nhiễm vào *E. coli*. Cosmid không chứa một gen nào của phagơ λ và nhờ vậy nó không hình thành nên vết tan sau khi lây nhiễm. Thay vào đó, tế bào vi khuẩn sau khi bị lây nhiễm được

nuôi cấy trên môi trường chứa agar và chất kháng sinh. Các khuẩn lạc có tính kháng kháng sinh mang cosmid tái tổ hợp sau đó được phân lập và nhân lên giống như phương pháp tách dòng plasmid. Các vector cosmid có ưu điểm là nó có thể mang những đoạn ADN ngoại lai có kích thước lớn. Nhờ các vector cosmid có kích thước nhỏ, điển hình khoảng 8kbp hoặc nhỏ hơn, và phần vỏ phagơ có thể mang phân tử ADN có kích thước đạt đến 52kbp, nên cosmid có thể được sử dụng để tách dòng các đoạn ADN ngoại lai có kích thước đạt đến 44kbp.

Nhiễm sắc thể nấm men nhân tạo (YAC)

Các vector loại này được phát triển trong thời gian gần đây và đại diện cho một hướng tách dòng gen mới mà trong đó vector tách dòng được sao chép trong tế bào chủ của sinh vật nhân chuẩn (eukaryote) giống như cách sao chép của nhiễm sắc thể của chính tế bào chủ đó. Nhiễm sắc thể nấm men nhân tạo YAC (*y*east *a*rtificial *c*hromosome) có tất cả các đặc tính của một nhiễm sắc thể cần cho quá trình tự sao chép của nó trong tế bào nấm men bao gồm một **vùng khởi đầu sao chép**, **một tâm động** để đảm bảo cho sự phân ly bình thường vào các tế bào con trong quá trình nguyên phân, và **các tâm mút** đảm bảo cho sự ổn định của nhiễm sắc thể (*Hình 5.8*). Nhiễm sắc thể nấm men nhân tạo có thể dùng để tách dòng các đoạn gen ngoại lai có kích thước rất lớn (đạt đến vài trăm kbp). Điều này có ý nghĩa đặc biệt quan trọng, vì nhờ vậy các dòng vector riêng biệt có thể mang các gen nguyên vẹn có nguồn gốc từ động vật có vú. Nhiễm sắc thể nấm men nhân tạo cũng có thể được sử dụng để xây dựng bản đồ hệ gen của người hoặc các loài sinh vật có hệ gen lớn thông qua việc xác định các dòng chứa các vùng nằm cận kề nhau trong hệ gen. Có hai loại vector có quan hệ với nhiễm sắc thể nhân tạo nấm men là nhiễm sắc thể nhân tạo vi khuẩn (BAC = *b*acterial *a*rtificial *c*hromosome) và nhiễm sắc thể nhân tạo P1 (PAC = P1 *a*rtificial *c*hromosome) được sử dụng giống như nhiễm sắc thể nấm men nhân tạo.



Hình 5.8. **Cấu trúc nhiễm sắc thể (NST) nhân tạo nấm men (YAC)**

Vector tách dòng thực vật

Phần lớn các vector tách dòng gen ở tế bào thực vật được xây dựng dựa trên cấu trúc của Ti plasmid vốn xuất hiện tự nhiên ở vi khuẩn trong đất có tên là *Agrobacterium tumefaciens*. Vi khuẩn này xâm nhập vào mô thực vật và gây nên sự phát triển của các tế bào "khối u" hay còn gọi là **nốt sần**. Trong quá trình lây nhiễm, một phần của Ti plasmid được gọi là vùng T-ADN kết hợp vào hệ gen của tế bào thực vật chủ. Các vector tách dòng dựa trên Ti plasmid lợi dụng khả năng kết hợp tự nhiên với hệ gen nhân của tế bào chủ của đoạn trình tự T-ADN để tiến hành biến nạp các gen ngoại lai có ích vào hệ gen thực vật; các gen được biến nạp thường liên quan đến các đặc tính có ích cho cây chủ, chẳng hạn như các gen kháng bệnh, gen tăng cường chất lượng, v.v...

Các ứng dụng của tách dòng ADN

Kỹ thuật tách dòng gen là một kỹ thuật sinh học quan trọng được sử dụng rộng rãi trong nhiều lĩnh vực nghiên cứu sinh học khác nhau, bao gồm các ví dụ dưới đây:

– **Xác định các gen gây bệnh.** Việc tách dòng gen giúp xác định các gen bị hỏng hoặc sai lệch, đột biến gây nên các bệnh di truyền, chẳng hạn như bệnh thiếu máu hồng cầu hình liềm, rối loạn dinh dưỡng cơ, các gen gây ung thư và gen ức chế khối u có vai trò quan trọng trong quá trình phát triển của khối u. Việc nghiên cứu đặc tính của các gen này giúp nâng cao hiểu biết về cơ chế phát sinh bệnh.

– **Lập bản đồ gen.** Việc xác định và làm rõ đặc tính các dòng gen nằm gần nhau trên nhiễm sắc thể được sử dụng để lập bản đồ gen. Nhờ ứng dụng này có thể xác định được vị trí tương đối của các gen trên nhiễm sắc thể.

– **Các protein tái tổ hợp.** Kỹ thuật tách dòng gen được sử dụng để sản xuất một lượng lớn các loại protein có giá trị trong y học, chẳng hạn như insulin trong điều trị bệnh đái tháo đường, protein yếu tố VIII trong điều trị chứng thiếu máu. Các gen mã hóa các loại protein hữu ích được tách dòng nhờ các vector biểu hiện và biến nạp vào một cơ thể chủ thích hợp chẳng hạn như nấm men để các gen tách dòng có thể được biểu hiện và sản xuất ra một lượng lớn protein theo yêu cầu.

– **Tạo ra các cơ thể biến đổi di truyền GMO (genetically *m*odified *o*rganism).** Người ta có thể sử dụng kỹ thuật tách dòng gen để khuếch đại, chuyển các gen ngoại lai mong muốn vào một cơ thể sinh vật nào đó để tạo ra những cơ thể (động vật, thực vật) mới mang các đặc tính di truyền mong muốn.

V.C PCR - PHẢN ỨNG CHUỖI TRÙNG HỢP

Nguyên tắc

Phản ứng chuỗi trùng hợp PCR (DNA *p*olymerase *c*hain *r*eaction) là một kỹ thuật hữu hiệu được sử dụng rộng rãi và có vai trò quan trọng trong hầu hết các nghiên cứu về gen hiện nay. ADN nhiễm sắc thể trong tế bào chứa hàng ngàn gen. Điều đó làm cho việc phân tích và phân lập một gen riêng lẻ trở nên khó khăn. Phản ứng PCR là công cụ cho phép chúng ta sao chép nhiều lần một trình tự ADN đặc thù, thông thường tương ứng với các gen hoặc từng phần của các gen đó, xuất phát từ mẫu ADN nhiễm sắc thể nhờ một phản ứng enzym đơn giản. Yêu cầu duy nhất của quá trình này là cần phải biết một số trình tự ở hai đầu của đoạn ADN cần khuếch đại. Đoạn ADN được quan tâm nghiên cứu được nhân lên (khuếch đại) hàng triệu lần và trở thành đoạn ADN chiếm ưu thế sau khi phản ứng PCR kết thúc. Lượng ADN của đoạn trình tự được quan tâm nghiên cứu lúc đó đủ lớn để có thể sử dụng trong các nghiên cứu chi tiết tiếp theo về gen hoặc sử dụng cho mục đích thao tác di truyền liên quan đến các đoạn gen được khuếch đại.

Thành phần phản ứng PCR

Đoạn ADN được khuếch đại nhờ phản ứng PCR diễn ra nhờ sự thay đổi nhiệt độ của phản ứng qua nhiều chu kỳ. Mỗi một phản ứng PCR gồm có 4 thành phần chính:

– **ADN mạch khuôn**. Phân tử ADN này chứa đoạn ADN cần khuếch đại. Phân tử ADN mạch khuôn thường là một hỗn hợp phức tạp của nhiều trình tự ADN khác nhau vì được chiết suất từ ADN nhiễm sắc thể. Tuy vậy, mọi phân tử ADN chứa đoạn trình tự đích đều có thể sử dụng làm khuôn cho phản ứng PCR. Phân tử ARN cũng có thể sử dụng làm mạch khuôn cho phản ứng PCR nhờ việc sử dụng enzym phiên mã ngược (reverse transcriptase) để chuyển trình tự ARN thành trình tự ADN.

– **Đoạn mồi oligonucleotid**. Một phản ứng PCR thường cần một cặp mồi oligonucleotid. Đây là những phân tử ADN mạch đơn ngắn (thường gồm khoảng 10 - 20 base nitơ) thu được nhờ quá trình tổng hợp hóa học. Các đoạn mồi được chọn sao cho trình tự các oligonucleotid của chúng bổ trợ với trình tự ở hai đầu đoạn ADN cần khuếch đại (trên mạch đối diện).

– **ADN polymerase**. Một số loại ADN polymerase được sử dụng trong phản ứng PCR. Những enzym này là các enzym bền ở nhiệt độ cao, có thể ổn định và hoạt động ở nhiệt độ tới 100°C. Enzym được sử dụng phổ biến nhất là *Taq* ADN polymerase có nguồn gốc từ *Thermus aquaticus*, một loại vi khuẩn được phát hiện ở suối nước nóng. Vai trò của ADN polymerase trong phản ứng PCR là xúc tác quá trình sao chép các phân tử ADN. Enzym này gắn vào chuỗi ADN mạch đơn và tổng hợp một mạch ADN mới có trình tự bổ trợ với mạch ADN làm khuôn (mạch gốc). Tuy vậy, enzym ADN polymerase đòi hỏi phải có một đoạn ADN mạch kép để có thể khởi đầu phản ứng trùng hợp phân tử ADN. Trong quá trình phản ứng PCR xảy ra, đoạn mồi oligonucleotid gắn vào hai đầu của đoạn ADN cần khuếch đại theo nguyên tắc bổ trợ và tạo ra các vùng ADN sợi kép ngắn. Nhờ vậy, các đoạn mồi giúp định hướng để enzym ADN polymerase tiến hành kéo dài chuỗi polynucleotid và chỉ sao chép đoạn trình tự ADN đích.

– **Các tiểu phần dNTP (deoxyribonucleotid triphosphate)**. Các đơn phân tử này tương ứng với bốn loại base nitơ có mặt trong phân tử ADN (adenine, guanine, thymine và cytosine), đồng thời chúng là cơ chất hoạt động của enzym ADN polymerase. Mỗi phản ứng PCR cần 4 loại dNTP (dATP, dGTP, dTTP, và dCTP) như các **tiểu phần cấu trúc** hình thành nên phân tử ADN nhờ hoạt động của enzym ADN polymerase.

Phản ứng PCR xảy ra như thế nào

Phản ứng PCR cho phép khuếch đại các đoạn trình tự ADN đích thông qua nhiều chu kỳ tổng hợp ADN lặp đi lặp lại nhiều lần (*Hình 5.9*). Mỗi một phân tử ADN đích được tổng hợp mới sau đó lại được sử dụng làm mạch khuôn cho chu kỳ tổng hợp tiếp theo. Kết quả là lượng ADN của đoạn trình tự đích tăng lên nhanh chóng sau mỗi chu kỳ và trở thành phân tử chiếm ưu thế sau khi phản ứng kết thúc. Trong các chu kỳ đầu tiên, quá trình tổng hợp ADN tăng lên theo hàm số mũ, nhưng ở các chu kỳ sau vì lượng ADN đích được sao chép tăng lên và các thành phần tham gia phản ứng được sử dụng hết, lượng ADN được khuếch đại tăng lên theo kiểu tuyến tính, sau đó không tăng lên nữa.

Mỗi một chu kỳ tổng hợp ADN liên quan đến 3 giai đoạn là **biến tính, gắn mồi, kéo dài chuỗi**, trong đó mỗi giai đoạn xảy ra ở một chế độ nhiệt khác nhau đến khi một phân tử ADN mới được tổng hợp hoàn chỉnh.

Giai đoạn biến tính:

Nhiệt độ phản ứng được nâng lên trên 90°C. Ở nhiệt độ này, cấu trúc ADN xoắn kép bị biến tính và tách ra thành hai mạch đơn riêng biệt và có thể dùng làm khuôn để tổng hợp nên phân tử ADN mới nhờ hoạt động của enzym ADN polymerase.

Giai đoạn gắn mồi:

Phản ứng được làm lạnh về nhiệt độ phù hợp cho việc gắn của các đoạn mồi vào phân tử ADN mạch đơn làm khuôn, nhưng không làm các mạch đơn làm khuôn liên kết trở lại với nhau thành mạch kép. Vì vậy, giai đoạn này gọi là **giai đoạn gắn mồi**. Nhiệt độ của giai đoạn này có thể thay đổi (35 - 60°C) tuỳ thuộc vào chủng loại và số lượng các nucleotid có trong đoạn mồi.

Giai đoạn kéo dài chuỗi:

Giai đoạn này diễn ra ở nhiệt độ mà enzym ADN polymerase có hoạt tính cao. Đối với *Taq* polymerase, nhiệt độ này là 72°C. Nhờ sự định hướng của đoạn mồi, enzym ADN polymerase tiến hành sao chép phân tử ADN theo nguyên tắc kéo dài chuỗi polynucleotid dựa trên một mạch ADN làm khuôn. Một phản ứng PCR điển hình thường gồm tổng cộng 20 - 40 chu kỳ tùy thuộc vào lượng trình tự ADN đích có trong phân tử ADN làm khuôn. Phản ứng PCR có thể sử dụng để khuếch đại các đoạn ADN có kích thước đến vài ngàn cặp base nitơ. Để có thể thực hiện một số lượng lớn các chu kỳ biến đổi nhiệt độ khác nhau, phản ứng PCR được tiến hành nhờ **máy biến nhiệt theo chu kỳ** (thermal cycler). Trong chu kỳ đầu tiên, các phân tử ADN được tổng hợp đến đầu tận cùng của chuỗi ADN gốc, đó là do không có sự ngăn cản enzym ADN polymerase tiến hành sao chép cho đến đầu tận cùng của mạch ADN gốc. Tuy vậy, trong các chu kỳ tiếp theo, phân tử ADN mới được tổng hợp có đầu tận cùng là trình tự của đoạn mồi được sử dụng làm khuôn, và vì vậy việc tổng hợp các phân tử ADN mới chỉ giới hạn trong đoạn trình tự được giới hạn bởi các đoạn mồi, hay nói cách khác chỉ đoạn trình tự ADN đích được tổng hợp.

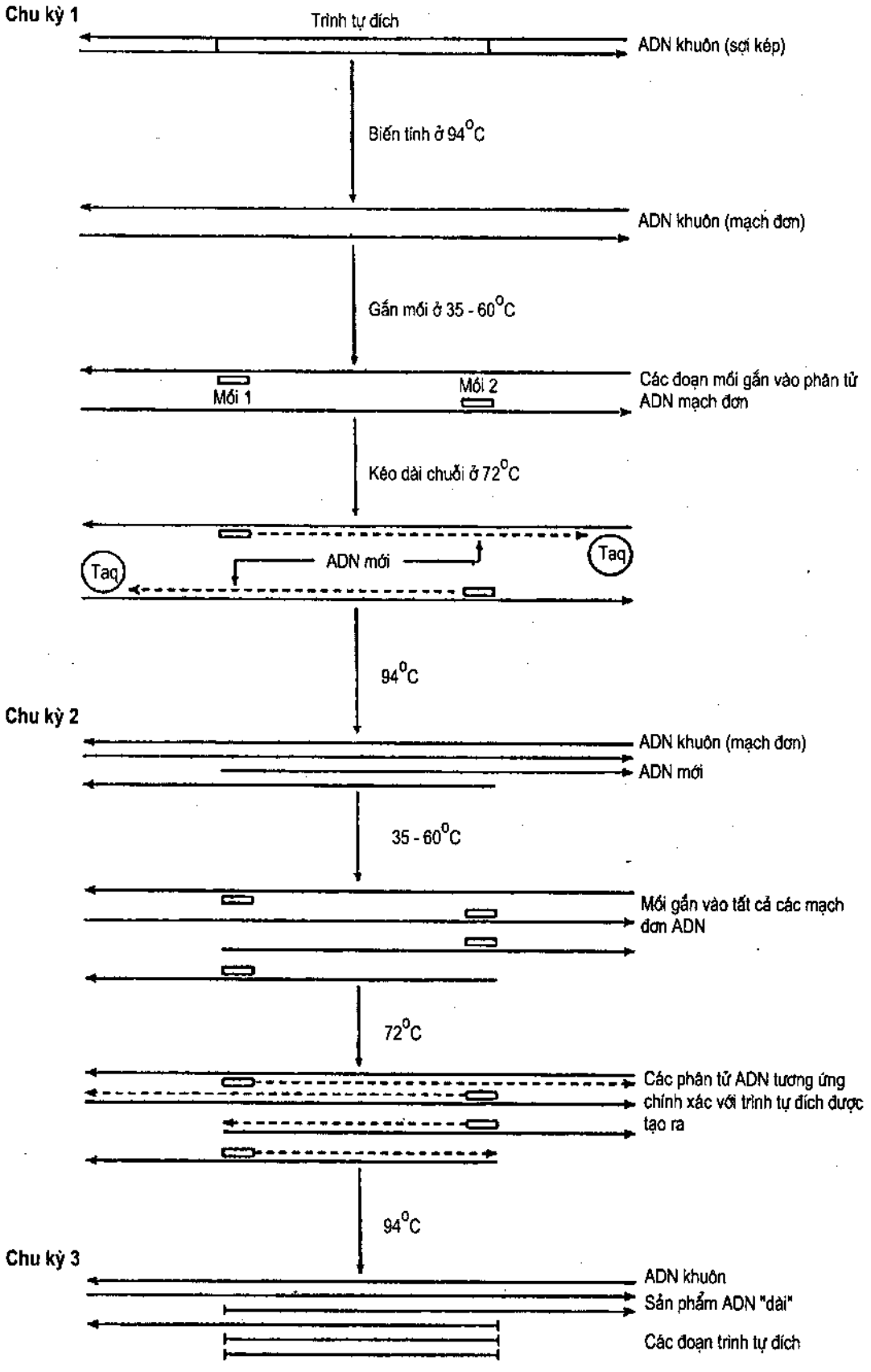
Các ứng dụng của phản ứng PCR

Kỹ thuật PCR được sử dụng rộng rãi như một phương tiện nghiên cứu giúp tăng cường các khả năng nghiên cứu về gen. Phần lớn các nghiên cứu về di truyền phân tử hiện nay đòi hỏi việc sử dụng kỹ thuật PCR như một kỹ thuật đầu tay chủ yếu, hoặc để phối hợp với các phương pháp nghiên cứu khác. Chẳng hạn như, sản phẩm PCR có thể được sử dụng để giải trình tự ADN, sử dụng làm mẫu dò trong các phương pháp thẩm tách Southern và Northern, và để tạo ra các dòng gen mới. Kỹ thuật PCR không chỉ được ứng dụng trong các lĩnh vực nghiên cứu sinh học và y học thực nghiệm, mà ngay cả trong các lĩnh vực sinh học đại cương như nhân chủng học và cổ sinh vật học.

Sau đây là một số lĩnh vực mà kỹ thuật PCR ngày nay được sử dụng rộng rãi và đã có những đóng góp đáng kể:

– **Nghiên cứu về các bệnh di truyền.** Các rối loạn bệnh lý gây ra do đột biến gen có thể được di truyền từ bố, mẹ sang con cái. Chẳng hạn như đối với bệnh máu khó đông và bệnh xơ nang, người ta có thể sử dụng kỹ thuật PCR để khuếch đại các trình tự gen để sàng lọc và tìm ra các đột biến gây bệnh. Thông tin thu được sẽ giúp cán bộ nghiên cứu hiểu biết chi tiết hơn về nguyên nhân của các dạng bệnh lý này, đồng thời có thể giúp xác định các cá thể mang gen gây bệnh.

– **Nghiên cứu về ung thư học.** PCR đã và đang được sử dụng rộng rãi trong các nghiên cứu tìm hiểu về vai trò của gen trong quá trình phát sinh các tế bào ung thư. Chẳng hạn như đột biến xảy ra ở các gen gây ung thư (oncogene), hoặc các gen ức chế khối u (tumor suppressor) có thể xác định được bằng việc khuếch đại ADN thu được từ các tế bào ung thư nhờ việc sử dụng kỹ thuật PCR. Điều này giúp cho các nhà nghiên cứu có hiểu biết chi tiết hơn về quá trình hình thành và phát triển của các tế bào ung thư.



Hình 5.9. **Phản ứng chuỗi trùng hợp - PCR**

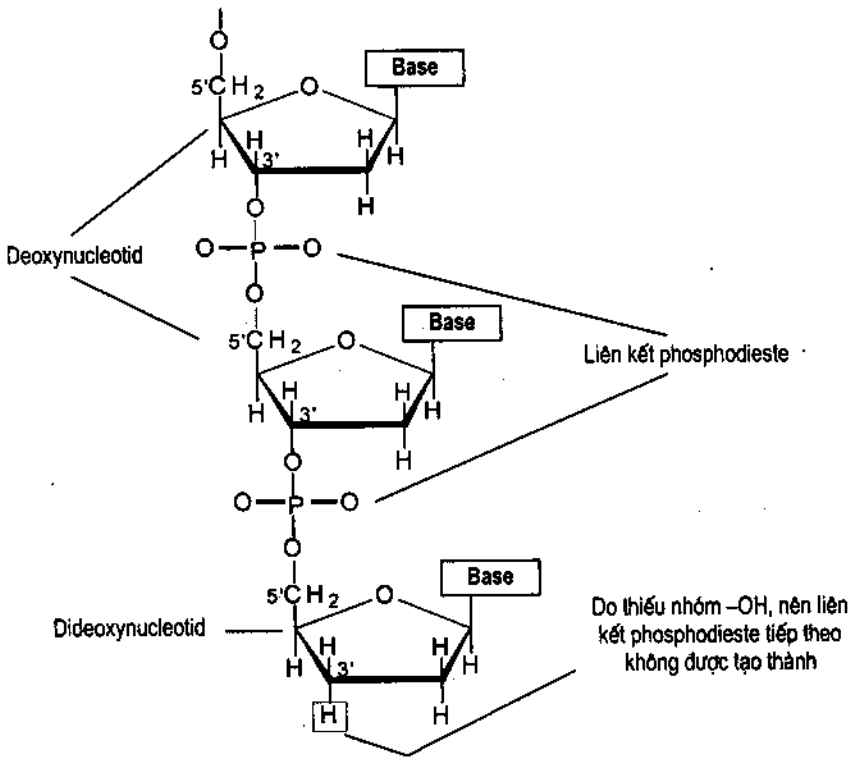
– **Nghiên cứu pháp y.** Nhờ khả năng có thể khuếch đại một đoạn ADN đặc trưng lên hàng triệu triệu lần, kỹ thuật PCR có thể được dùng để xác định các cá thể nhờ vào mẫu ADN của cá thể đó. Chẳng hạn, các cá thể có thể được xác định dựa trên mẫu ADN còn sót lại ở hiện trường các vụ án mạng. Ngoài ra, kỹ thuật PCR cũng tạo điều kiện cho các nghiên cứu xác định các "dòng vận động" của gen trong các nghiên cứu nhân chủng học và nguồn gốc các loài, trong đó có các chủng tộc người khác nhau.

– **Công nghệ sinh học.** PCR giữ một vai trò quan trọng trong việc sản xuất ra các protein tái tổ hợp, chẳng hạn như insulin hay các hormon sinh trưởng vốn được sử dụng làm thuốc điều trị, cũng như trong việc nghiên cứu và phát triển các loại vaccin tái tổ hợp, ví dụ như vaccin tái tổ hợp chống virus viêm gan B, v.v...

V.D GIẢI MÃ TRÌNH TỰ ADN

Nguyên lý giải mã trình tự ADN

Có hai nguyên lý giải mã trình tự các nucleotid của phân tử ADN đã được phát triển, đó là phương pháp kết thúc chuỗi dideoxy (ddNTP) của Sanger và phương pháp biến tính hóa học của Maxam và Gilbert. Phương pháp kết thúc chuỗi (sử dụng ddNTP) hiện nay được sử dụng thay thế cho phương pháp hóa học do kỹ thuật này hiệu quả và dễ thực hiện hơn.



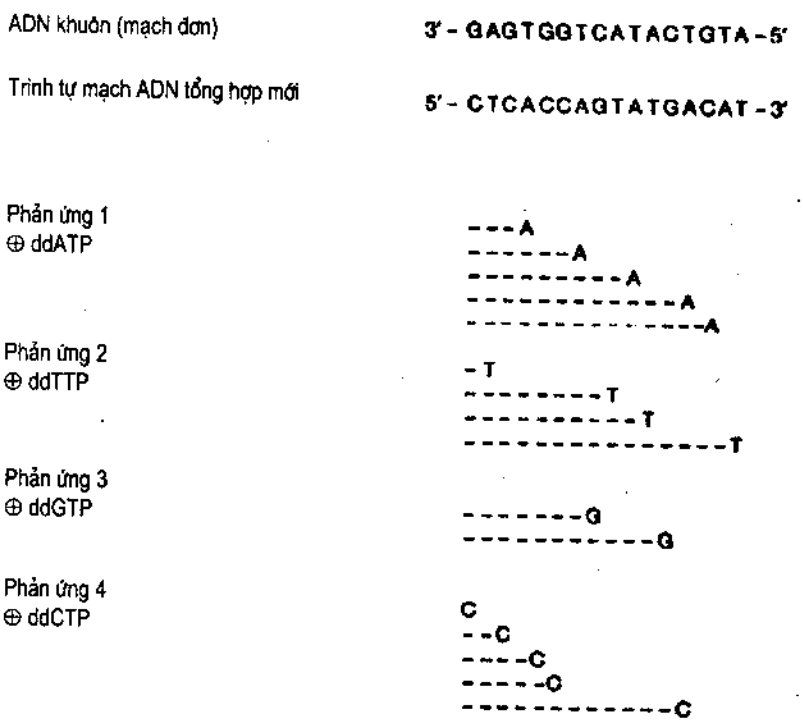
Hình 5.10. **Sự kết thúc tổng hợp ADN bởi dideoxynucleotid (ddNTP)**

Nguyên lý của phương pháp kết thúc chuỗi dideoxy là dựa trên việc phân tử ADN được sao chép nhờ hoạt động của enzym ADN polymerase. Việc sử dụng các phân tử nucleotid

triphosphate cải tiến thành dạng dideoxynucleotid triphosphate (ddNTP) làm dừng phản ứng tổng hợp ADN một cách ngẫu nhiên tại vị trí tương ứng với một trong bốn loại base nitơ trên mạch ADN làm khuôn. Kết quả là một dãy các phân tử ADN sai khác nhau về chiều dài được tạo ra. Các phân tử này sau đó sẽ được phân tách nhờ phương pháp điện di trên gel. Trình tự của các base nitơ trên mạch ADN làm khuôn được xác định bằng việc xác định loại base nitơ có mặt ở đầu tận cùng của mỗi phân tử ADN được tổng hợp kể từ phân tử ngắn nhất cho đến phân tử dài nhất.

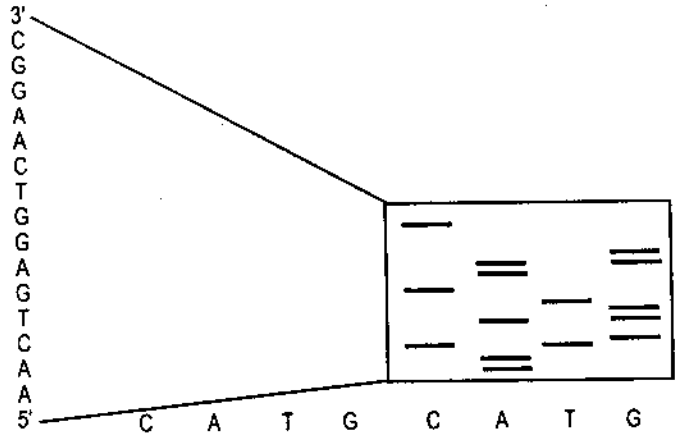
Phương pháp kết thúc chuỗi trùng hợp nhờ dideoxynucleotid

Phân tử ADN được giải mã trình tự gọi là **mạch khuôn**. Phân tử này phải được làm tinh sạch và đồng nhất, nghĩa là nó chỉ chứa các phân tử ADN có trình tự giống hệt nhau. Các vector plasmid tinh sạch chứa trình tự ADN được tách dòng hoặc các phân tử ADN được tạo ra nhờ phản ứng PCR thường được sử dụng làm mạch khuôn. Một yêu cầu đối với quá trình giải mã trình tự sử dụng phương pháp kết thúc chuỗi trùng hợp là phân tử ADN phải ở trạng thái mạch đơn là dạng có thể sao chép bởi enzym ADN polymerase. Trước đây, việc này có thể thực hiện nhờ việc tách dòng phân tử ADN bằng vector phagơ M13. Khi vector phagơ tái tổ hợp được dùng để lây nhiễm vi khuẩn *E. coli*, nó tạo ra các bản sao của trình tự ADN được tách dòng ở trạng thái mạch đơn và vì vậy có thể sử dụng để làm mạch khuôn cho quá trình giải mã trình tự. Gần đây, người ta thường sử dụng phương pháp đơn giản hơn nhiều là gây biến tính ADN bằng việc xử lý nhiệt hoặc bằng dung dịch kiềm để thu phân tử ADN ở trạng thái mạch đơn sử dụng cho việc giải mã trình tự, qua đó không phải tiến hành việc tách dòng phân tử. Có một số loại ADN polymerase được sử dụng cho mục đích giải mã trình tự, bao gồm một phần của enzym ADN polymerase I của *E. coli* gọi là **đoạn Klenow**, đây một enzym ADN polymerase được cải biến di truyền có nguồn gốc từ phagơ T7 là **Sequencase** và ***Taq* ADN polymerase** vốn được sử dụng trong phản ứng PCR.



Hình 5.11. **Giải mã trình tự theo nguyên tắc kết thúc chuỗi bởi dideoxynucleotid (ddNTP)**

Tất cả các enzym ADN polymerase yêu cầu phải có một vùng ADN sợi kép ngắn ở đầu để có thể khởi đầu việc xúc tác tổng hợp phân tử ADN dựa trên trình tự nucleotid trên mạch đơn làm khuôn. Vì vậy, phản ứng giải mã trình tự phải được bổ sung bằng các phân tử ADN mạch đơn ngắn gọi là các **đoạn mồi oligonucleotid** được tạo ra bằng phương pháp tổng hợp hóa học. Các đoạn mồi được chọn sao cho trình tự các nucleotid của nó bổ trợ với trình tự ADN mạch khuôn. Nhờ vậy, các đoạn mồi có thể gắn vào mạch ADN làm khuôn tạo nên một vùng ngắn ADN cấu trúc mạch kép, đó là điểm phản ứng giải mã trình tự bắt đầu diễn ra. Để giải mã trình tự một phân tử ADN, người ta chuẩn bị 4 phản ứng enzym riêng biệt. Mỗi phản ứng bao gồm: phân tử ADN làm khuôn ở trạng thái mạch đơn, ADN polymerase, đoạn mồi, hỗn hợp bốn loại deoxyribonucleotid triphosphate (dNTP) là đơn phân tử cấu thành nên phân tử ADN trong phản ứng trùng hợp, và một loại nucleotid cải tiến được gọi là **dideoxyribonucleotid triphosphate** (ddNTP). Có bốn loại ddNTP khác nhau tương ứng với bốn loại base nitơ trong phân tử ADN. Các phân tử ddNTP khác với dNTP ở chỗ không có nhóm hydroxyl (-OH) ở vị trí C-3' của phân tử đường pentose. Phân tử ddNTP có thể liên kết được vào chuỗi ADN đang được trùng hợp nhờ hoạt tính của enzym ADN polymerase, nhưng vì thiếu nhóm -OH ở vị trí C-3', liên kết phosphodiester không được hình thành với nucleotid ở vị trí kế tiếp. Vì vậy, quá trình kéo dài chuỗi ADN bị dừng lại (*Hình 5.10*). Có thể nói ddNTP là yếu tố kìm hãm phản ứng tổng hợp ADN. Vì mỗi một trong bốn phản ứng giải mã trình tự chứa một loại ddNTP khác nhau, nên phụ thuộc vào loại ddNTP nào có mặt trong phản ứng, quá trình tổng hợp ADN sẽ kết thúc ở vị trí của nucleotid đó. Vào mọi thời điểm trong quá trình tổng hợp ADN, enzym ADN polymerase có thể xúc tác phản ứng liên kết của dNTP vào chuỗi polynuleotid đang được tổng hợp, lúc đó phản ứng kéo dài chuỗi tiếp tục xảy ra, hoặc nó có thể xúc tác để ddNTP liên kết vào chuỗi đang được mở rộng, nhưng trong trường hợp đó phản ứng tổng hợp ADN bị dừng lại. Kết quả cuối cùng là mỗi một trong số 4 phản ứng giải mã trình tự sẽ tạo ra một dãy các phân tử ADN có kích thước khác nhau, và mỗi phân tử sẽ có trình tự được kết thúc tại vị trí có mặt của base nitơ tương ứng trên phân tử ADN mạch khuôn (*Hình 5.11*). Ví dụ như trong phản ứng giải mã trình tự chứa ddATP, một dãy các phân tử ADN sẽ được tổng hợp trong đó mỗi phân tử sẽ kết thúc ở vị trí của A (adenin) tương ứng với vị trí của T trên mạch khuôn.



Hình 5.12. Thang giải mã trình tự ADN

Khi phản ứng giải mã trình tự kết thúc, các phân tử ADN mới được tổng hợp được phân tách nhờ phương pháp điện di trên gel polyacrylamid với sản phẩm của 4 phản ứng giải mã được chạy trên các giếng gần nhau cùng lúc. Bằng việc bổ sung thêm dNTP được đánh dấu phóng xạ nhờ các đồng vị lưu huỳnh và phospho, các phân tử ADN được tổng hợp sẽ có hoạt tính phóng xạ, vì vậy có thể phát hiện được khi đưa bản gel polyacrylamide lên tấm phim chụp tia X. Phương pháp này gọi là phương pháp **phóng xạ tự chụp**. Sau khi xử lý phim tia X, một dãy các băng điện di hiện lên trên bề mặt của phim tương ứng với 4 dải điện di chạy từ 4 giếng của 4 phản ứng giải mã trình tự khác nhau. Trình tự của mạch ADN khuôn nhờ vậy có thể xác định bằng việc đọc từ các băng có kích thước nhỏ nhất lần lượt đến các băng có kích thước lớn hơn liền kề (*Hình 5.12*).

Giải mã trình tự ADN tự động

Mặc dù nguyên lý cơ bản của phương pháp giải mã trình tự không thay đổi, nhưng kỹ thuật này được phát triển lên thành một kỹ thuật tự động, bao gồm một hệ thống bán tự động trong đó công việc điện di sản phẩm ADN khuếch đại và việc phát hiện cũng như phân tích các phản ứng giải mã trình tự được thực hiện bởi một hệ thống thiết bị do máy tính điều khiển. Với thiết kế như vậy, bốn loại ddNTP được đánh dấu bằng các loại thuốc nhuộm màu khác nhau thông qua liên kết cộng hóa trị. Chỉ cần thực hiện một phản ứng giải mã trình tự duy nhất và sản phẩm tạo ra là một dãy các phân tử ADN mà phân tử này chỉ dài hơn phân tử kia một nucleotid duy nhất được đánh dấu bằng thuốc nhuộm có màu tương ứng nằm ở vị trí cuối cùng của chuỗi polynuleotid. Các sản phẩm ADN được khuếch đại sau đó được điện di trên gel polyacrylamid. Nucleotid cuối cùng nằm trên chuỗi phát ra ánh sáng huỳnh quang khi được chiếu bằng nguồn sáng laser và được ghi nhận bằng một thiết bị dò có sẵn trong hệ thống. Nhờ vào bước sóng của ánh sáng phát ra, thiết bị dò có thể xác định được nucleotid cuối cùng trong mỗi chuỗi polynucleotid, đồng thời xây dựng một "bức tranh toàn cảnh" về thành phần và trình tự các nucleotid của phân tử ADN được khuếch đại.

Phương pháp giải mã trình tự ADN tự động có nhiều ưu thế vượt trội so với các phương pháp truyền thống hoặc phương pháp hóa học (thực hiện bằng tay). Mỗi một phản ứng giải mã trình tự đối với một mẫu chỉ đòi hỏi một ống nghiệm đựng mẫu và chạy trên một dải điện di duy nhất, trong khi số lượng này trong phương pháp truyền thống là bốn. Đặc tính này kết hợp với khả năng phát hiện và phân tích mẫu tự động giúp làm tăng tốc độ giải mã trình tự và dung lượng dữ liệu có thể quản lý khai thác một cách rõ rệt.

Phân tích trình tự ADN được giải mã

Giải mã trình tự ADN là một trong những kỹ thuật nền tảng của lĩnh vực di truyền học phân tử và ngày càng được sử dụng rộng rãi trong các nghiên cứu. Điểm quan trọng của kỹ thuật này là khả năng xác định thông tin liên quan đến thành phần của ADN. Trong những năm gần đây, lượng thông tin liên quan đến trình tự ADN được xác định ở các loài sinh vật khác nhau đã tăng lên nhanh chóng, đến mức những thông tin này phải được quản lý dưới dạng các ngân hàng dữ liệu, trong đó có các ngân hàng dữ liệu **EMBL** ở Châu Âu và **Genbank** ở Mỹ. Các ngân hàng dữ liệu này chứa các trình tự của các gen và hệ gen người cũng như của nhiều loài sinh vật khác nhau được các nhóm nghiên cứu khác nhau trên thế giới tìm ra. Những thông tin về trình tự ADN trong các ngân

hàng dữ liệu như vậy hiện nay có thể dễ dàng truy cập và tải về miễn phí thông qua mạng Internet.

Việc giải mã trình tự ADN giúp tạo ra một lượng số liệu khổng lồ. Để giúp cho việc phân tích các trình tự được dễ dàng hơn, người ta sử dụng một số phần mềm chuyên dụng. Chẳng hạn như phần mềm GCG của Đại học Wisconsin, v.v... cho phép phân tích nhanh chóng các đặc tính quan trọng của các trình tự ADN, như vị trí cắt của enzym giới hạn, mã bộ ba mở đầu, các mã bộ ba kết thúc, khung đọc, phần giao nhau giữa exon - intron và các trình tự gen điều hòa. Các gen này cũng cho phép so sánh các trình tự mới được phát hiện với các trình tự đã có trong ngân hàng dữ liệu và qua đó cho biết mức độ giống nhau, khác nhau và mối quan hệ của các gen này.

Việc phát triển các hệ thống giải mã trình tự ADN tự động đã giúp việc xác định trình tự ADN của toàn bộ hệ gen một số loài sinh vật quan trọng trở nên dễ dàng hơn. Cho đến nay toàn bộ hệ gen của *E. coli* ($4,6 \times 10^6$bp), nấm men *S. cereviseae* ($2,3 \times 10^7$bp), người ($6,4 \times 10^9$bp) và một số loài sinh vật khác đã được xác định. Ngoài ra, dự án giải mã trình tự hệ gen cơ bản của người đến nay cũng đã hoàn thành.

Phần VI
CÁC ỨNG DỤNG CỦA DI TRUYỀN HỌC

Phần VI gồm các mục:

VI.A Các bệnh di truyền

Các bệnh di truyền
Hình thức di truyền
Sai hỏng đơn gen
Các đột biến trong sai hỏng đơn gen
Phát hiện cá thể mang bệnh

VI.B Di truyền học và ung thư

Ung thư là bệnh di truyền
Gen gây ung thư
Chuyển đoạn nhiễm sắc thể đặc trưng ung thư
Gen ức chế khối u
Các dạng ung thư di truyền

VI.C Phép thử độc tính gen

Phép thử độc tính gen
Phép thử Ames
Phép thử gây đột biến dòng tế bào
Phép thử di truyền học tế bào

VI.D Liệu pháp gen

Tổng quan
Vector chuyển gen
Thiếu năng enzym adenin deaminase
Bệnh xơ nang
Bệnh ung thư

VI.E Dự án hệ gen người

Cơ sở khoa học
Bản đồ di truyền
Bản đồ vật lý
Phương pháp lai huỳnh quang *in situ*
Sắp xếp vị trí các gen trên bản đồ

VI.F Di truyền học và khoa học hình sự

Sự tương quan đặc trưng cá thể
So sánh mẫu protein
So sánh ADN
RFLP: dấu vân tay di truyền
VNTR: hồ sơ di truyền
Phương pháp phân tích dữ liệu

VI.G Kỹ thuật di truyền và công nghệ sinh học

Kỹ thuật di truyền và công nghệ sinh học
Sự biểu hiện của các protein tái tổ hợp
Hệ thống biểu hiện protein ở vi khuẩn
Hệ thống biểu hiện ở sinh vật nhân chuẩn
Kỹ thuật di truyền thực vật
Động vật chuyển gen

VI. A CÁC BỆNH DI TRUYỀN

Các bệnh di truyền

Có một nhóm đa dạng các bệnh lý và rối loạn gây ra do các đột biến gen và sự thay đổi bất thường của nhiễm sắc thể. Các rối loạn có bản chất di truyền và có thể chia làm 3 nhóm chính:

Các sai hỏng đơn gen:

Các sai hỏng đơn gen còn được gọi là **các rối loạn di truyền Mendel** (Mendelian disorders), **các rối loạn đơn gen** (monogenic disorders), hay **các rối loạn đơn locut** (single locus disorders). Đây là một nhóm các dạng bệnh lý gây ra do sự có mặt của một gen đột biến trong cơ thể bị bệnh. Đột biến gen làm thay đổi thông tin mã hóa của gen đó và, hoặc dẫn đến việc tạo ra phân tử protein bị sai hỏng về chức năng, hoặc thậm chí ức chế hoàn toàn sự tổng hợp protein mà gen đó mã hóa. Sự thiếu hụt protein do đột biến gen gây nên sự biểu hiện của các trạng thái bệnh lý. Đột biến gen có thể được di truyền giữa các thế hệ (từ bố, mẹ sang con, cháu) hoặc xuất hiện một cách tự phát (*de novo*) trong tế bào sinh dục (tinh trùng hoặc trứng) trong cơ thể bố hoặc mẹ, và sau thụ tinh, đứa trẻ hình thành mang đột biến trong mọi tế bào.

Các rối loạn nhiễm sắc thể:

Có các dạng bệnh lý gây ra do sự mất đi hoặc thêm vào một hoặc một số nhiễm sắc thể, hay do sự thay đổi cấu trúc của nhiễm sắc thể. Phần lớn các rối loạn bất thường về nhiễm sắc thể xuất hiện ngay trong các tế bào sinh dục của cơ thể bố hoặc mẹ, nhưng cũng có những trường hợp gây ra do di truyền từ thế hệ trước. Các dạng bất thường về số lượng nhiễm sắc thể (biến dị số lượng nhiễm sắc thể) có thể biểu hiện bằng sự tăng lên số lượng bộ nhiễm sắc thể đơn bội (**hiện tượng đa bội thể**), hoặc do sự thêm và/hoặc mất đi của từng nhiễm sắc thể riêng lẻ (**hiện tượng lệch bội**). Các dạng bất thường về cấu trúc nhiễm sắc thể có thể gây ra do sự đứt gẫy nhiễm sắc thể liên quan đến các hiện tượng mất đoạn, lặp đoạn hoặc đảo đoạn nhiễm sắc thể.

Các rối loạn đa nhân tố:

Đây là một nhóm gồm nhiều bệnh phổ biến, ví dụ như đái tháo đường, các bệnh mạch vành và phần lớn các dị tật bẩm sinh. Các bệnh này gây ra do ảnh hưởng của nhiều gen theo các cơ chế bệnh lý phức tạp cho đến nay chưa được hiểu biết đầy đủ, nhưng được biết có liên quan đến sự tương tác của nhiều gen với nhau, hoặc giữa các gen với các yếu tố môi trường.

Trong khoảng 20 năm qua, nhờ sự phát triển của công nghệ ADN tái tổ hợp, đã có nhiều phát hiện mới mang tính bước ngoặt liên quan đến các bệnh lý do rối loạn đơn gen gây ra. Thông tin được nêu trong phần dưới đây liên quan đến một số rối loạn bệnh lý như vậy.

Hình thức di truyền

Các rối loạn di truyền đơn gen được truyền từ thế hệ bố, mẹ sang thế hệ con, cháu. Có ba hình thức di truyền phổ biến: **di truyền trội trên nhiễm sắc thể thường, di truyền lặn trên nhiễm sắc thể thường** và **di truyền liên kết nhiễm sắc thể X** (*Bảng 6.1*).

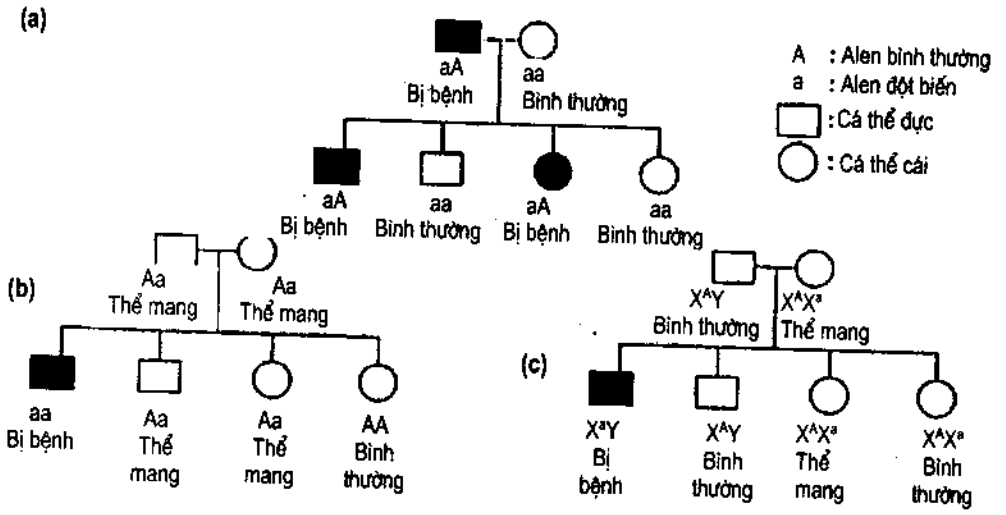
Trong trường hợp bệnh di truyền do alen trội nằm trên nhiễm sắc thể thường quy định, việc truyền một alen gây bệnh từ bố hoặc mẹ sang con là đủ để cá thể con biểu hiện bệnh. Các

cá thể bị bệnh có một alen bình thường và một alen đột biến gây bệnh được gọi là các **thể dị hợp tử**. Các cá thể này có nguy cơ truyền cho 50% số con alen đột biến và biểu hiện bệnh (*Hình 6.1a*).

Trong trường hợp bệnh di truyền do alen lặn nằm trên nhiễm sắc thể thường quy định, cá thể biểu hiện bệnh phải mang đủ một cặp alen đột biến gây bệnh, một bắt nguồn từ bố, một từ mẹ. Cá thể biểu hiện bệnh trong trường hợp này gọi là **cá thể đồng hợp tử** về alen đột biến. Các cá thể dị hợp tử với một alen gây bệnh không biểu hiện bệnh nhưng có khả năng truyền alen gây bệnh sang 50% số cá thể con. Đối với trường hợp cả bố và mẹ là các cá thể dị hợp tử mang alen lặn gây bệnh trên nhiễm sắc thể thường, một phần tư số con biểu hiện bệnh, một phần tư bình thường và một nửa số cá thể con là thể mang alen gây bệnh nhưng không biểu hiện bệnh (*Hình 6.1b*).

Trong trường hợp bệnh di truyền liên kết với nhiễm sắc thể X, gen đột biến gây bệnh chỉ xuất hiện trên nhiễm sắc thể X. Do con đực chỉ có một nhiễm sắc thể X duy nhất, việc truyền alen đột biến sang cá thể con giới đực là đủ để cá thể này biểu hiện bệnh. Các cá thể đực biểu hiện bệnh gọi là các cá thể dị giao tử. Các con cái có hai nhiễm sắc thể X, vì vậy thường không biểu hiện bệnh do phần lớn các gen đột biến gây bệnh nằm trên nhiễm sắc thể X là các alen lặn. Đối với các con cái là cá thể mang gen gây bệnh nhưng không biểu hiện bệnh, 50% con đực ở thế hệ con có nguy cơ bị bệnh và 50% con cái là thể mang gen gây bệnh nhưng không biểu hiện bệnh (*Hình 6.1c*).

Để có sự cân bằng giữa con đực và con cái về lượng sản phẩm do gen nằm trên nhiễm sắc thể X mã hóa, trong tự nhiên có hiện tượng một trong hai nhiễm sắc thể X trong tế bào con cái bị bất hoạt. Quá trình này được gọi là **hiện tượng Lyon hóa (giả thiết Lyon)** và thường diễn ra trong quá trình phát triển của phôi. Trong mỗi tế bào, nhiễm sắc thể X bị bất hoạt được "chọn" một cách ngẫu nhiên. Tuy vậy, một số con cái là thể mang gen gây bệnh nằm trên nhiễm sắc thể X có thể biểu hiện bệnh ở mức độ nhẹ do sự bất hoạt của nhiễm sắc thể X bình thường.



Hình 6.1. (a) Alen trội trên NST thường. Sự di truyền của một alen đột biến duy nhất (A) đều dẫn đến sự biểu hiện của bệnh. (b) Alen lặn trên NST thường. Các cá thể bị bệnh phải mang hai alen đột biến (aa). Các cá thể dị hợp tử (Aa) là các thể mang. (c) Alen liên kết NST X, đối với các thể mang là cái, 50% số cá thể con giới đực bị bệnh, và 50% số cá thể con giới cái là thể mang.

Sai hỏng đơn gen

Có nhiều bệnh di truyền do gen đơn quy định xuất hiện ở người (*Bảng 6.1*). Các bệnh này có các đặc điểm biểu hiện đa dạng khác nhau và hậu quả đối với các cá thể bị bệnh cũng khác nhau tùy thuộc vào mức độ quan trọng của gen bị đột biến và bản chất của loại đột biến xuất hiện. Một số bệnh, như bệnh máu khó đông, gây nên các triệu chứng bệnh có thể điều trị được, nhưng các bệnh khác, chẳng hạn như hội chứng múa giật Huntington, đến nay chưa có biện pháp điều trị triệt để và người bệnh thường chết khi còn trẻ.

Các bệnh di truyền do đơn gen quy định thường xuất hiện với tần số tương đối thấp nằm trong khoảng giữa 0,03 đến 0,5 trường hợp trong 1000 em bé sơ sinh (xem *Bảng 6.1*). Tần số các rối loạn di truyền này thường khác nhau trong các chủng tộc người khác nhau. Chẳng hạn, tần số bệnh nhân bị xơ nang là cao nhất ở các nước Bắc Âu, bệnh thiếu máu hồng cầu hình liềm xảy ra với tần số cao nhất ở Châu Phi và bệnh β-thalassemia phổ biến hơn cả trong các quần thể Châu Á. Đối với các bệnh di truyền đơn gen phổ biến nhất ở người đến nay đã xác định và tách dòng được gen gây bệnh, đồng thời xác định được các đột biến gây bệnh.

Bảng 6.1. **Một số bệnh lý di truyền đơn gen**

Bệnh lý	*Tần số trên 1000 trẻ sơ sinh*	*Hình thức di truyền*	*Gen đột biến*	*Đặc điểm*
Máu khó đông dạng A	0,1	Liên kết NST X	Nhân tố VIII	Chảy máu bất thường
Máu khó đông dạng B	0,03	Liên kết NST X	Nhân tố IX	Chảy máu bất thường
Loạn dưỡng cơ Duchene	0,3	Liên kết NST X	Dystrophin	Hao mòn cơ
Loạn dưỡng cơ Becker	0,05	Liên kết NST X	Dystrophin	Hao mòn cơ
Hội chứng NST X yếu	0,5	Liên kết NST X	FMR1	Chậm phát triển trí tuệ
Bệnh múa giật Huntington	0,5	Trội, trên NST thường	Hungtingtin	Chứng tâm thần phân liệt
U sơ thần kinh	0,4	Trội, trên NST thường	NF-1,2	Ung thư
Hội chứng thalassemia	0,05	Lặn, trên NST thường	Các gen globin	Thiếu máu
Thiếu máu hồng cầu hình liềm	0,1	Lặn, trên NST thường	β - globin	Thiếu máu; Thiếu máu cục bộ
Phenylketo niệu	0,1	Lặn, trên NST thường	Phenylalanine-hydroxylase	Không có khả năng chuyển hóa phenylalanin
Xơ nang	0,4	Lặn, trên NST thường	CFTR	Bệnh hỏng phổi tích lũy và các triệu chứng khác

Các đột biến trong sai hỏng đơn gen

Có thể chia các loại đột biến tạo ra các alen gây bệnh thành hai loại chính: **các đột biến điểm** liên quan đến sự thay đổi của một base nitơ duy nhất và **các đột biến lớn** liên quan đến sự thay đổi trình tự ADN với kích thước lớn hơn. Đối với mỗi loại bệnh, có thể có vài dạng đột biến khác nhau. Ngoài ra, các cá thể bị bệnh cũng có thể cùng lúc mang các gen đột biến khác nhau. Ví dụ, có khoảng 20% trường hợp bị bệnh máu khó đông dạng A do kết quả của đột biến lớn gây ra. Các trường hợp còn lại là do các dạng đột biến điểm mà đến nay các nhà nghiên cứu đã tìm ra và mô tả 250 kiểu đột biến khác nhau.

Các đột biến điểm:

Các đột biến điểm gây nên các bệnh di truyền có thể chia thành một số kiểu sau:

(1) Các đột biến sai nghĩa, nhầm nghĩa (misense mutations). Đây là những thay đổi của các nucleotid trên phân tử ADN gây nên sự thay đổi bộ ba mã hóa cho một amino acid dẫn đến sự thay thế bởi một loại amino acid khác trên phân tử protein. Các đột biến sai nghĩa gây nên những hậu quả khác nhau đối với phân tử protein được mã hóa. Do hiện tượng thoái hóa của mã di truyền, những thay đổi liên quan đến vị trí base thứ ba trong bộ ba mã hóa thường không có ảnh hưởng đến phân tử protein. Ngoài ra, nhiều sự thay đổi thành phần base nitơ dẫn đến sự thay thế của amino acid có đặc tính tương tự có thể không làm thay đổi chức năng và hoạt tính của phân tử protein. Chẳng hạn như đột biến ở bộ ba mã hóa CTT thành ATT làm thay thế amino acid kỵ nước là leucin bằng isoleucin cũng là một amino acid kỵ nước khác. Tuy vậy có nhiều ví dụ cho thấy, các đột biến sai nghĩa làm thay đổi rõ rệt chức năng của phân tử protein được mã hóa và vì vậy gây nên các bệnh di truyền. Trong số này có thể kể đến đột biến thay thế A bằng T trong gen mã hóa β-globin, một trong các chuỗi polypeptid hình thành nên phân tử hemoglobin. Đột biến này làm thay đổi bộ ba số sáu của gen, từ GAG mã hóa cho acid glutamic thành GTG mã hóa cho valin. Đột biến này gây nên bệnh thiếu máu hồng cầu hình liềm do các tế bào hồng cầu bị biến dạng thành hình liềm do thay đổi sự kết dính của các phân tử hemoglobin. Các tế bào hồng cầu hình liềm có tuổi thọ ngắn gây nên hiện tượng thiếu máu và nằm trong các mao mạch làm giảm khả năng cung cấp máu tới các cơ quan (chứng thiếu máu cục bộ).

(2) Các đột biến vô nghĩa. Đây là những thay đổi của các nucleotid trên phân tử ADN làm chuyển một mã bộ ba mã hóa amino acid thành một mã bộ ba kết thúc, vì vậy quá trình phiên mã sẽ kết thúc sớm hơn bình thường và dẫn đến sự hình thành phân tử protein có kích thước ngắn hơn. Các đột biến vô nghĩa thường gây hậu quả nghiêm trọng đối với phân tử protein được mã hóa, đặc biệt khi nó xuất hiện gần đầu 5' của gen. Nhiều bệnh di truyền khác nhau đã được xác định có liên quan đến các đột biến vô nghĩa. Ví dụ như đột biến C thành T ở bộ ba số 39 trong gen mã hóa β-globin làm thay đổi mã bộ ba bình thường CAG quy định glutamin thành TAG là một bộ ba mã kết thúc. Đột biến này gây nên sự kết thúc phiên mã sớm của phân tử mARN mã hóa cho β-globin, dẫn đến sự thiếu hụt một chuỗi polypeptid β và gây nên dạng bệnh lý gọi là β-thalassemia với triệu chứng bệnh thiếu máu do phân tử hemoglobin bình thường không được tạo thành.

(3) Các đột biến dịch khung. Những đột biến này xảy ra do sự thêm vào hay mất đi của một hay một số base nitơ làm thay đổi khung đọc và một tập hợp các bộ ba mã hóa mới được hình thành kể từ điểm đột biến xảy ra. Đột biến dịch khung cũng thường gây nên hậu quả nghiêm trọng đối với phân tử protein được mã hóa, đặc biệt khi đột biến xuất hiện gần đầu 5' của gen. Nhiều bệnh lý được mô tả liên quan đến đột biến dịch khung. Chẳng hạn đột biến dịch khung đã được tìm thấy là nguyên nhân gây nên bệnh máu khó đông ở nhiều bệnh nhân mắc căn bệnh này. Trong đó bao gồm các trường hợp do mất đi 4 base nitơ gây nên sự thay đổi khung đọc từ bộ ba mã hóa thứ 50 và một đột biến thêm 10 base làm thay đổi khung đọc từ bộ ba mã hóa thứ 38. Cả hai kiểu đột biến này đều gây triệu chứng bệnh nghiêm trọng.

(4) Đột biến vị trí cắt intron. Đây là những đột biến làm thay đổi trình tự tín hiệu ở gần đầu 3' hoặc 5' của các đoạn intron dẫn đến việc cắt intron sai trong quá trình hoàn thiện phân tử mARN ở sinh vật nhân chuẩn. Các đột biến kiểu này cũng có thể xảy ra bên trong intron tạo nên điểm cắt intron mới và vì vậy cũng dẫn đến sự cắt sai trình tự intron. Một loạt các đột biến vị trí cắt intron được tìm thấy liên quan đến đột biến gen β-globin làm thiếu hoàn toàn các chuỗi β-globin trong các cơ thể đồng hợp tử và gây bệnh β-thalassemia.

(5) Đột biến trình tự gen điều hòa. Các đột biến này xảy ra tương đối hiếm và ảnh hưởng đến việc điều hòa hoạt động của gen, thường hoặc làm giảm hoặc làm tăng mức độ biểu hiện của gen. Một đột biến như vậy đã được xác định trong trình tự chỉ huy của gen mã hóa protein đông máu (là protein yếu tố IX) cũng là một nguyên nhân gây nên bệnh máu khó đông. Các cá thể mang đột biến này không tạo được protein yếu tố IX và bị chảy máu một cách bất thường. Thông thường, triệu chứng bệnh thường mất đi sau tuổi dậy thì nhờ hormon steroid kích thích sự biểu hiện của gen này.

Các đột biến lớn:

Có nhiều bệnh lý gây ra do các đột biến liên quan đến một trình tự dài các nucleotid trên phân tử ADN. Phần lớn các đột biến này có ảnh hưởng nghiêm trọng đến chức năng của gen và gây bệnh nặng.

(1) Các đột biến mất đoạn. Sự mất đi của gen có thể biểu hiện với mức độ kích thước khác nhau, từ một vài base nitơ đến toàn bộ gen, thậm chí nhiều gen cùng lúc. Sự mất đi hoàn toàn của các gen mã hóa β-globin gây nên bệnh β-thalassemia (bệnh mất khả năng sản xuất hemoglobin bình thường). Ví dụ như, sự mất một phần gen mã hóa dystrophin gây nên bệnh mòn cơ, bệnh loạn dưỡng cơ; hay sự mất đi một bộ ba mã hóa duy nhất trong gen tổng hợp protein điều hòa độ dẫn xuyên màng trong bệnh xơ nang CFTR (*c*ystic *f*ibrosis *t*ransmembrane *c*onductance *r*egulator) là nguyên nhân gây bệnh gặp phải ở 70% số bệnh nhân bị bệnh xơ nang.

(2) Các đột biến thêm đoạn. Nhiều đột biến thêm đoạn đã được ghi nhận. Ví dụ như một trường hợp một bệnh nhân bị máu khó đông dạng A hiếm gặp có nguyên nhân gây bệnh là do sự thêm vào gen mã hóa yếu tố VIII một trình tự lặp lại gọi là yếu tố LINE.

(3) Các đột biến thay thế đoạn gen. Cũng có nhiều đột biến thay thế đoạn gen gây nên bệnh di truyền đã được ghi nhận. Ví dụ như một đột biến gây bệnh máu khó đông dạng A xảy ra do sự tái tổ hợp giữa các trình tự nằm trong vùng intron thứ 22 của gen mã hóa yếu tố VIII và các trình tự lặp lại kép dọc theo nhiễm sắc thể X. Do một lỗi xảy ra trong quá trình tái tổ

hợp, gen mã hóa yếu tố VIII bị cắt thành 2 mảnh tách biệt nhau bởi hàng triệu cặp base nitơ, làm mất hoàn toàn chức năng của gen này.

(4) Các đột biến lặp lại bộ ba nucleotid. Một dạng đột biến gen hiếm gặp liên quan đến các trình tự lặp lại từng bộ ba nucleotid kém bền vững. Trong quá trình giảm phân xảy ra hiện tượng số lượng bản sao các trình tự lặp lại từng bộ ba nucleotid tăng lên trong các tế bào sinh dục dẫn đến sự biểu hiện của bệnh trong các thế hệ sau. Cơ chế dẫn đến hiện tượng lặp lại nhiều lần của các trình tự nucleotid và nguyên lý gây bệnh cho đến nay chưa được biết rõ. Sự tăng lên số lượng các trình tự lặp lại tìm thấy liên quan đến một số bệnh di truyền, bao gồm bệnh múa giật Hungtington.

Phát hiện cá thể mang bệnh

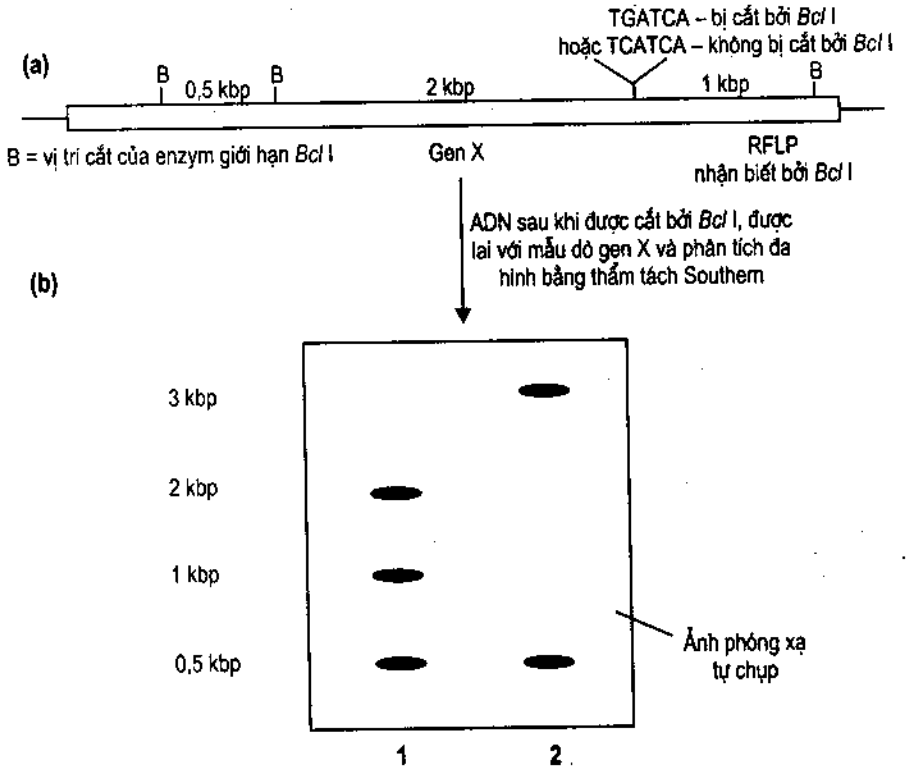
Trong các dạng bệnh lý di truyền do gen lặn quy định, các cá thể dị hợp tử mang alen đột biến gây bệnh nhưng không biểu hiện triệu chứng bệnh lý và có thể truyền alen gây bệnh đó sang thế hệ con. Nếu cả hai bố, mẹ đều là những cá thể mang dị hợp tử mang alen lặn gây bệnh thì 1/4 số con sinh ra có khả năng mang cả hai alen gây bệnh và biểu hiện bệnh. Vì vậy, việc xác định được các cá thể mang alen gây bệnh sẽ là cơ sở cho việc **tư vấn di truyền** trong hôn nhân.

Các cá thể mang alen gây bệnh có thể xác định được bằng phương pháp dò gen, theo đó các alen đột biến được xác định và hình thức di truyền của chúng giữa các thế hệ trong một dòng họ được theo dõi. Phương pháp dò gen sử dụng những biến dị tự nhiên trong các trình tự ADN của hệ gen người. Việc so sánh một trình tự gen giữa một nhóm cá thể thường biểu hiện sự biến đổi trong trình tự ADN theo kiểu thay thế của các base nitơ. Chẳng hạn trong một trình tự ADN nhất định, ở một số cá thể là nucleotid C, trong khi ở một số cá thể khác là T. Những biến đổi như vậy gọi là **tính đa hình** và là kết quả của những đột biến xảy ra trong phân tử ADN. Tính đa hình xảy ra rất phổ biến và xuất hiện với tần số khoảng 1/(100-150) base nitơ dọc suốt hệ gen. Khi chúng xảy ra trong các trình tự mã hóa của gen, chúng sẽ gây nên các bệnh di truyền và trở thành đối tượng của quá trình chọn lọc. Khi chúng xuất hiện trong các trình tự không mã hóa, chúng thường không gây nên những hậu quả nghiêm trọng cho cơ thể sinh vật nhưng có xu hướng tích lũy theo thời gian. Một số dạng đa hình biểu hiện trong các trình tự đích của enzym giới hạn. Phụ thuộc vào các nucleotid được thay thế, một trình tự cắt của một enzym giới hạn nào đó có thể xuất hiện hoặc mất đi. Tính đa hình xuất hiện ở vị trí cắt của enzym giới hạn gọi là **tính đa hình độ dài đoạn phân cắt giới hạn RFLP** (*r*estriction *f*ragment *l*ength *p*olymorphism). RFLP là một công cụ hiệu quả khi được sử dụng làm dấu chuẩn của các gen và có thể xác định được nhờ phương pháp thẩm tách Southern. Nếu một phân tử ADN hệ gen được cắt bởi một enzym giới hạn, kích thước của các đoạn ADN được tạo ra sẽ biến đổi tùy thuộc vào việc xuất hiện hay vắng mặt của vị trí cắt giới hạn. Kiểu hình các đoạn cắt được xác định bởi phương pháp thẩm tách Southern đối với ADN của các cá thể khác nhau là khác nhau phụ thuộc vào sự có mặt hay vắng mặt của các trình tự giới hạn (*Hình 6.2*). Kiểu hình các đoạn cắt giới hạn cũng cho phép phân biệt hai alen khác nhau cùng có mặt trong một cơ thể. Thông qua việc phân tích ADN của một cá thể bị bệnh, người ta có thể xác định được tính đa hình có liên quan đến một alen đột biến hay không, cũng như có thể xác định được hình thức di truyền của alen gây bệnh. Bằng việc phân tích

kiểu hình đoạn phân cắt giới hạn của các thành viên trong một gia đình, có thể xác định được các cá thể con không biểu hiện bệnh mang alen bình thường, hoặc có mang alen gây bệnh, qua đó phân biệt được các thể mang alen gây bệnh.

Kỹ thuật dò gen cũng được sử dụng để thực hiện phép **chẩn đoán tiền sản** để xác định trạng thái của bào thai trước khi sinh. Một mẫu ADN của bào thai có thể được phân tích theo dấu chuẩn RFLP sử dụng kỹ thuật thẩm tách Southern để xác định xem bào thai đó là bình thường, hay là thể dị hợp tử, hoặc đồng hợp tử về alen gây bệnh.

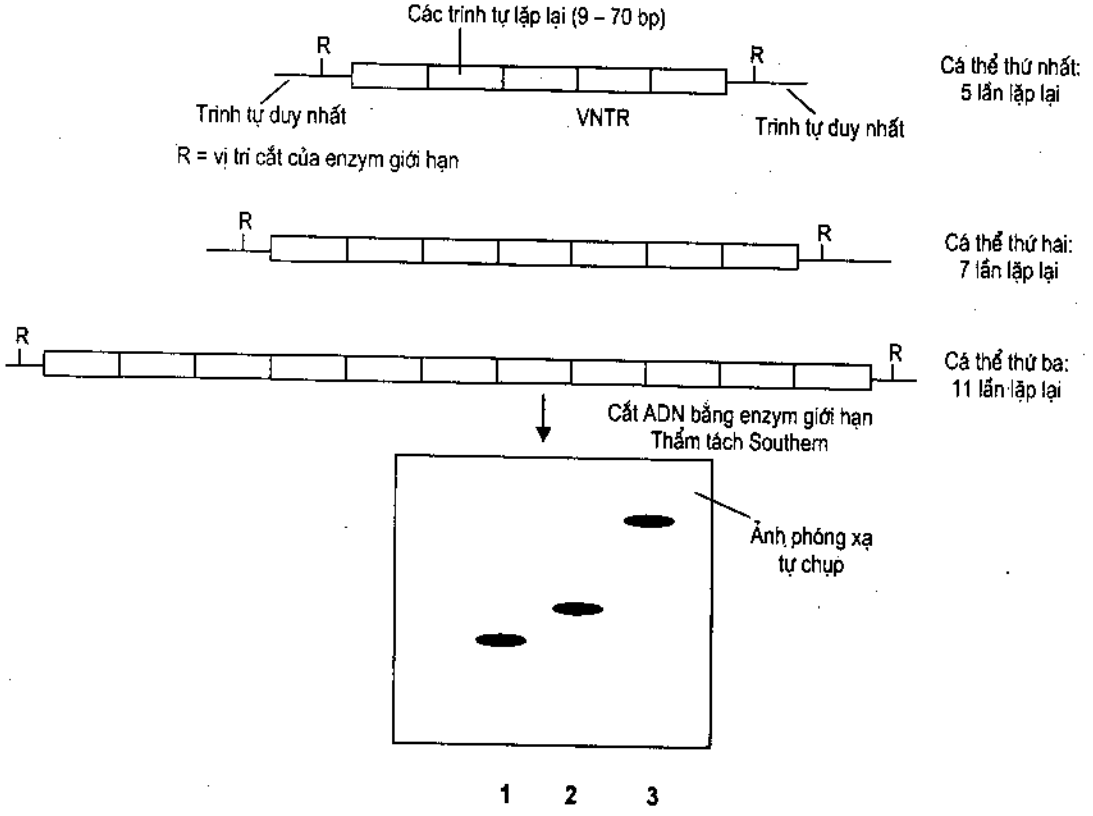
Việc sử dụng kỹ thuật RFLP nhằm xác định các cá thể mang các bệnh di truyền dù sao cũng bị hạn chế do tính biến dị ADN tại các điểm phân cắt giới hạn thường là thấp. Chỉ có hai trình tự có thể xuất hiện ở một vị trí RFLP, vì vậy trong nhiều trường hợp các alen đột biến và alen bình thường có cùng trình tự ở vị trí đó nên không phân biệt được. Trong những trường hợp này, việc xác định cá thể mang alen gây bệnh không thực hiện được. Hạn chế này của kỹ thuật RFLP có thể được khắc phục nhờ việc sử dụng các trình tự lặp lại liên tục vì những trình tự này có mức độ biến dị cao hơn.



Hình 6.2. **(a) Gen X có một dấu chuẩn RFLP được nhận biết bởi enzym giới hạn *Bcl* I. (b) Trong cá thể đồng hợp tử, kỹ thuật thẩm tách Southern tạo 3 phân đoạn (đường chạy điện di 1: có kích thước tương ứng là 2, 1, và 0,5kbp) khi có mặt vị trí cắt của *Bcl* và 2 phân đoạn (đường chạy điện di 2, có kích thước tương ứng là 3 và 0,5kbp) khi vắng mặt vị trí cắt của enzym giới hạn *Bcl* I.**

Bên cạnh các trình tự xuất hiện một bản sao duy nhất, hệ gen người còn chứa một tỷ lệ lớn các trình tự lặp lại và xuất hiện thành từng cụm trong đó phổ biến là các trình tự ngắn lặp đi lặp lại nhiều lần (xem mục II.D). Số lần lặp lại của các trình tự này có mức độ biến động lớn giữa các cá thể. Các trình tự lặp lại đầu tiên được xác định là các trình tự ngắn có chiều dài

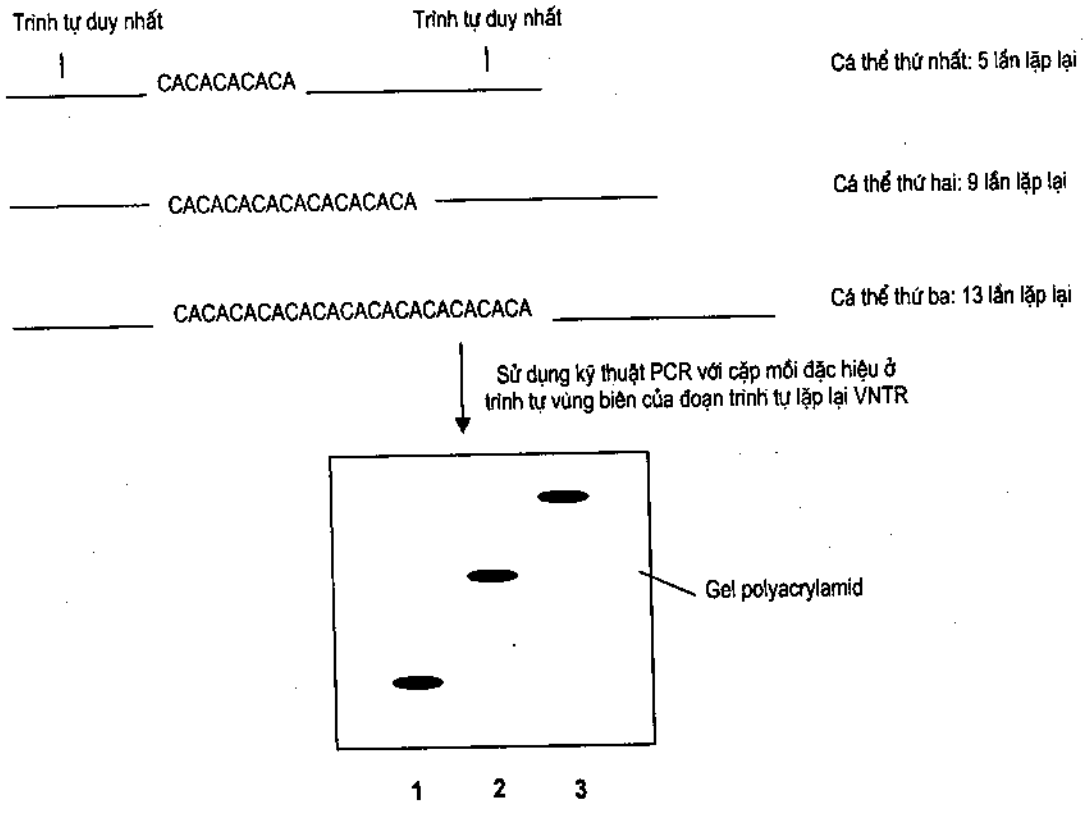
biến động giữa 9 và 70bp và xếp thành một chuỗi liên tục từ đầu đến cuối cụm trình tự lặp lại. Các cụm trình tự ADN gồm các trình tự ngắn lặp lại như vậy gọi là cụm **trình tự lặp lại liên tiếp với số lượng biến đổi VNTR** (*v*ariable *n*umber *t*andem *r*epeats) hay còn gọi là **trình tự ADN tiểu vệ tinh** (minisatelite DNA). Số lần lặp lại trong mỗi khối VNTR khác nhau giữa các cá thể giao động từ một vài bản sao đến hàng trăm bản sao khác nhau. Bằng việc cắt phân tử ADN nhờ enzym giới hạn ở hai đầu (vùng biên) của cụm VNTR, các đoạn ADN có chiều dài khác nhau được phát hiện nhờ kỹ thuật thẩm tách Southern (*Hình 6.3*). Một loại trình tự lặp lại phổ biến khác được gọi là các trình tự lặp lại CA, hay còn gọi là các trình tự ADN lặp lại **vi vệ tinh** (microsatellite DNA). Trong trường hợp này, trình tự lặp lại là cặp nucleotid CA thường xuất hiện thành cụm có số lần lặp lại giao động trong khoảng 10 - 60 lần dọc suốt hệ gen. Người ta sử dụng kỹ thuật PCR để khuếch đại đoạn trình tự lặp lại CA để thu các phân tử ADN có chiều dài khác nhau rồi phân tách chúng trên gel điện di (*Hình 6.4*). Cả hai loại trình tự lặp là CA và VNTR có thể sử dụng như những gen đánh dấu và dùng để xác định các cá thể mang alen gây bệnh di truyền giống như cách người ta sử dụng kỹ thuật RFLP.



Hình 6.3. **Chỉ thị đa hình VNTR có thể phát hiện được bằng kỹ thuật thẩm tách Southern và sử dụng để phân biệt các cá thể (sơ đồ trên đây minh họa ở cá thể đồng hợp tử).**

Một nhược điểm của kỹ thuật dò gen là cần phải phân tích cả các thành viên khác trong gia đình của người mang alen. Điều này là một khó khăn vì có thể không có đầy đủ các thành viên trong gia đình có mặt hoặc sẵn sàng tham gia vào nghiên cứu. Một phương pháp khác có thể dò gen gây bệnh là xác định các cá thể mang alen gây bệnh bằng việc xác định các đột

biến gây bệnh. Phương pháp này không đòi hỏi việc phân tích các thành viên khác nhau trong một gia đình, và nhờ vậy quá trình phân tích đơn giản hơn. Tuy vậy, trong phương pháp này trước tiên cần phải xác định đột biến có liên quan, và đôi khi việc này gặp khó khăn nếu gen có liên quan gây bệnh có kích thước lớn.



Hình 6.4. **Đoạn trình tự lặp lại CA với số lượng biến động có thể được khuếch đại nhờ phản ứng PCR và phân tích trên gel polyacrylamid. Các cá thể khác nhau có thể phân biệt được nhờ độ dài đoạn ADN được khuếch đại (sơ đồ này minh hoạ cá thể đồng hợp tử).**

VI.B DI TRUYỀN HỌC VÀ UNG THƯ

Ung thư là bệnh di truyền

Có nhiều bằng chứng cho thấy ung thư là một bệnh gây ra do sự thay đổi của các gen. Sau đây là một số dẫn chứng:

– Quá trình nguyên phân thường diễn ra kém chính xác trong các tế bào ung thư, dẫn đến sự thay đổi khác nhau về số lượng nhiễm sắc thể trong các tế bào của cùng một khối u.

– Các nhiễm sắc thể trong các khối u thường có sự sắp xếp lại về mặt cấu trúc. Mặc dù phần lớn những sự sắp xếp này xảy ra một cách ngẫu nhiên, nhưng cũng tồn tại một số dạng ung thư là kết quả của sự sắp xếp lại nhiễm sắc thể vốn có tính đặc thù cao.

– Phần lớn các chất gây đột biến cũng là các **chất gây ung thư**.

– Hầu như tất cả các tế bào ung thư khi nuôi cấy đều hình thành nên các dòng tế bào ung thư mới.

– Khuynh hướng mắc một hay một số dạng ung thư có biểu hiện di truyền trong một số dòng họ.

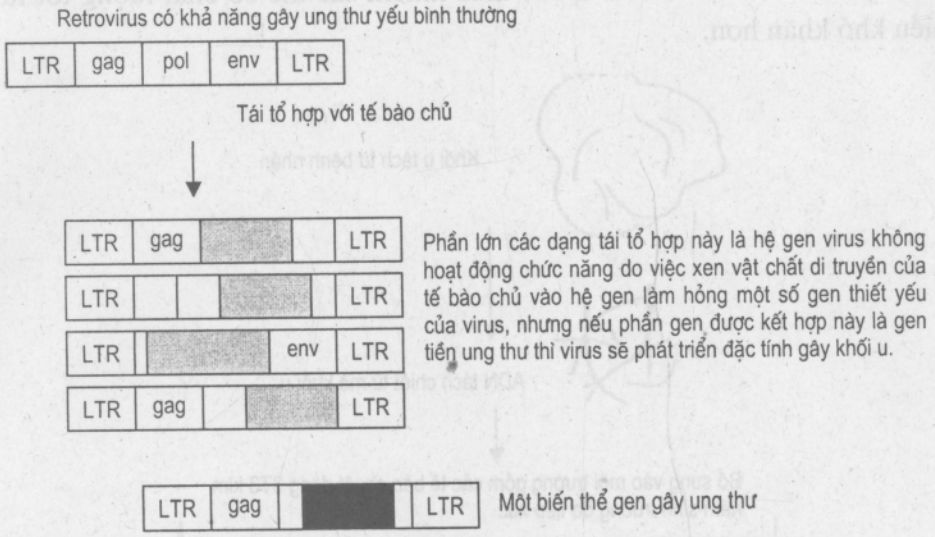
Gen gây ung thư

Các gen này lần đầu tiên được phát hiện trong một nhóm virus động vật gây ung thư, nhóm retrovirus. Nhóm virus này có hệ gen bản chất là ARN nằm trong các hạt virus, nhưng sau khi lây nhiễm vào tế bào chủ, hệ gen này được chuyển thành ADN nhờ hoạt động của enzym **reverse transcriptase**, và kết hợp vào hệ gen ADN nhân của tế bào chủ. Các dòng retrovirus gây ung thư khác các dòng không gây ung thư ở chỗ chúng có một gen khác. Gen này được gọi là **gen gây ung thư** (oncogene). Các trình tự ADN có quan hệ gần gũi với các gen gây ung thư ở virus cũng được tìm thấy trong ADN từ các cá thể động vật khoẻ mạnh bình thường. Các loại retrovirus gây ung thư là kết quả của quá trình tái tổ hợp di truyền giữa các dòng virus không gây ung thư và các gen trong tế bào chủ (*Hình 6.5*). Để phân biệt giữa các bản sao oncogene của virus với các bản sao của gen trong tế bào chủ, các bản sao của gen trong tế bào chủ được gọi là **các gen gây ung thư sơ khai** (proto-oncogene). Các gen gây ung thư của virus khác các gen gây ung thư sơ khai là nó có thể bị đột biến và không mang các trình tự intron. Khi có mặt trong retrovirus, nó chịu sự điều khiển của một trình tự gen tăng cường mạnh, đó là đoạn trình tự lặp lại dài ở đầu tận cùng LTR (*l*ong *t*erminal *r*epeat), làm tăng mức độ phiên mã của gen. Quá trình đột biến và/hoặc tăng cường mức độ biểu hiện của gen được gọi là quá trình hoạt hóa gen gây ung thư.

Các gen gây ung thư cũng tìm thấy từ các khối u không có sự xâm nhập của virus. Để phát hiện được điều đó, người ta tách chiết ADN từ các tế bào khối u và chuyển vào một dòng tế bào chuột 3T3. Các tế bào này sinh trưởng trong môi trường nuôi cấy tế bào và phát triển bên trên bề mặt của đĩa nuôi cấy. Sau khi chúng phân chia và tăng lên đủ để phủ kín bề mặt đĩa, các tế bào dừng phân chia... Quá trình này gọi là sự **ức chế do tiếp xúc**. Việc có mặt một gen gây ung thư được hoạt hóa vượt qua "rào cản" của quá trình ức chế này, và tế bào được biến nạp ADN chứa một gen gây ung thư tiếp tục phân chia để hình thành nên ổ tế bào gồm nhiều lớp. Các ổ tế bào này được phân lập, nuôi cấy và tách chiết gen gây ung thư (các trình tự ADN người có thể được xác định trên nền tế bào chuột nuôi cấy nhờ sự có mặt của trình tự lặp lại *Alu* đặc thù ở linh trưởng). Các gen gây ung thư được hoạt hóa không liên quan đến hoạt động của virus được gọi là *c - onc* (cellular oncogene) nhằm phân biệt với các gen gây ung thư của virus, gọi tắt là *v - onc* (viral oncogene).

Các gen gây ung thư thường được đặt tên dựa trên chủng virus mà chúng được xác định lần đầu tiên hoặc theo tên loại khối u chúng được phân lập đầu tiên bằng phép thử lây nhiễm. Vì vậy, *v-myc* là gen gây ung thư được phân lập đầu tiên ở virus gây bệnh u tủy ở gia cầm, và *c-neu* được xác định lần đầu tiên từ ADN tách chiết từ một khối u trung tính ở chuột. Các loại protein được mã hóa bởi các gen gây ung thư thường có liên quan chặt chẽ với **các con đường truyền tín hiệu**. Đây là những hệ thống gồm nhiều bước có khả năng nhận tín hiệu sinh trưởng từ bề mặt tế bào và truyền tín hiệu đó vào trong nhân tế bào, nơi mà sự phiên mã của các gen liên quan đến sự khởi đầu quá trình phân chia tế bào. Một số gen gây ung thư liên quan đến sự điều hòa **quá trình chết theo chương trình của tế bào** (apotosis). Đây là một quá trình mà các tế bào được lập chương trình tự động chết nhằm phản ứng lại một số tín hiệu gây hại cho cơ thể như sự lây nhiễm virus hay khi ADN bị hư hỏng.

Sự hoạt hóa của một gen gây ung thư dẫn đến các hậu quả là hoặc làm tăng sinh các mô và tế bào cấu trúc, hoặc ngăn cản quá trình chết của tế bào. Cả hai xu hướng này đều dẫn đến sự tăng lên số lượng tế bào. Trong một số trường hợp, sự hoạt hóa các gen gây ung thư xuất hiện thông qua quá trình khuếch đại gen. Người ta đã tìm thấy hàng ngàn bản sao của một gen gây ung thư xuất hiện trong mỗi tế bào khối u. Một ví dụ về trường hợp này là sự khuếch đại của gen *c-myc* trong các tế bào khối u của nguyên bào thần kinh.



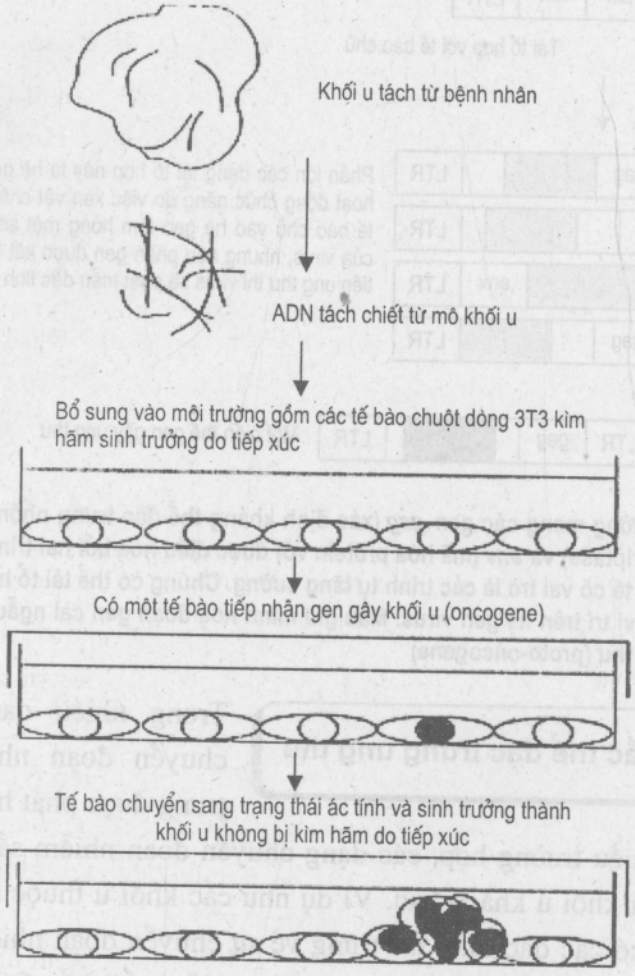
Hình 6.5. **Các retrovirus bình thường mang các gen *gag* (xác định kháng thể đặc trưng nhóm), *pol* (mã hóa enzym phiên mã ngược - reverse transcriptase) và *env* (mã hóa protein vỏ) được điều hòa bởi hai trình tự LTR (đoạn lặp dài ở đầu, *l*ong *t*erminal *r*epeat) thực tế có vai trò là các trình tự tăng cường. Chúng có thể tái tổ hợp với gen tế bào chủ. Tái tổ hợp có thể xảy ra ở nhiều vị trí trên hệ gen virus. Màu ghi minh hoạ đoạn gen cài ngẫu nhiên, màu đen minh hoạ đoạn gen cài là gen tiền ung thư (proto-oncogene)**

Chuyển đoạn nhiễm sắc thể đặc trưng ung thư

Trong nhiều dạng ung thư, các chuyển đoạn nhiễm sắc thể đặc trưng được phát hiện thấy trong hầu như mọi khối u. Trong nhiều trường hợp, các dạng chuyển đoạn nhiễm sắc thể này được sử dụng để phân biệt các loại khối u khác nhau. Ví dụ như các khối u thuộc hệ máu như các tế bào bạch cầu và lympho có các đặc tính đặc trưng về sự chuyển đoạn nhiễm sắc thể. Các tế bào lympho Burkitt có các chuyển đoạn đặc trưng giữa nhiễm sắc thể số 8 và các nhiễm sắc thể số 14, 2 hoặc 22. Trong đó, chuyển đoạn nhiễm sắc thể xảy ra giữa các nhiễm sắc thể số 8 và 14, t(8;14)(q24; q32) được quan sát thấy phổ biến nhất. Chuyển đoạn này mang gen gây ung thư sơ khai *myc* trên nhiễm sắc thể số 8 đến nằm ngay cạnh vùng gen mã hóa chuỗi nặng của globulin miễn dịch trên nhiễm sắc thể số 14. Gen *myc* được hoạt hóa bằng việc được chuyển nằm cạnh vùng tăng cường của gen quy định chuỗi nặng globulin miễn dịch. Gen này sau đó được phiên mã ở tốc độ tăng lên dẫn đến làm tăng lượng protein mà nó mã hóa và kích thích sự phân chia tế bào. Sự chuyển đoạn gen *myc* đến nhiễm sắc thể số 2 và 22 làm gen này nằm cạnh các vùng mã hóa chuỗi nhẹ của globulin miễn dịch.

Hiện tượng chuyển đoạn cũng có thể là nguyên nhân làm cho hai đầu tận cùng 5' và 3' của hai gen khác nhau kết hợp để tạo thành một gen mới. Gen mới này có thể là gen gây ung thư. Ví dụ về kiểu chuyển đoạn này là trường hợp gọi là **nhiễm sắc thể philadelphia** trong khối u của tế bào bạch cầu tủy, mà sự chuyển đoạn t(9; 12)(q34; q11) dẫn đến sự kết hợp hai gen ung thư sơ khai *bcr* và *abl* thành một gen gây ung thư mới.

Các quá trình chuyển đoạn đặc trưng ít được phát hiện hơn đối với các khối u cứng. Điều này có thể phản ánh việc chuẩn bị các mẫu nhiễm sắc thể có chất lượng tốt từ các khối u này gặp nhiều khó khăn hơn.



Hình 6.6. **Phát hiện gen gây khối u hoạt động từ ADN tách chiết từ mô khối u bằng phương pháp biến nạp**

Gen ức chế khối u

Trong số các gen thuộc loại này, cả hai alen phải bị bất hoạt để sự hình thành khối u có thể diễn ra. Các đột biến là lặn. Sự bất hoạt thường xảy ra nhất là do đột biến điểm ở một alen, và sự mất của alen còn lại do hiện tượng mất đoạn nhiễm sắc thể, mặc dù đôi khi hiện tượng đột biến ở cả hai alen cũng được quan sát thấy ở tế bào ung thư. Sự bất hoạt của các gen ức chế khối u xuất hiện trong các tế bào sinh dưỡng trong quá trình phát triển của một khối u. Gen ức chế ung thư được nghiên cứu kỹ nhất là *p53*. Gen này được đặt tên như vậy là vì nó mã hóa cho một phân tử phosphoprotein nhân có khối lượng 53kDa có chức năng là một yếu tố phiên mã. Gen này

bị đột biến trong hơn 40% các loại khối u ở người. *p53* liên quan đến nhiều quá trình hoạt động tế bào khác nhau. Nó có vai trò chính trong thời điểm chuyển từ pha G1 sang pha S của chu kỳ tế bào. Đây là một **điểm kiểm tra tế bào** nhằm xác định có "dừng" sự chuyển giữa các pha của tế bào trong chu kỳ tế bào hay không. Ở dạng tự nhiên, *p53* ngăn cản tế bào chuyển vào pha S, nhưng khi nó bị phosphoryl hóa (gắn thêm nhóm phosphat), nó mất chức năng này. Đây là một phần của quá trình điều hòa chu trình tế bào bình thường. Đột biến ở gen *p53* làm mất khả năng "bắt giữ" tế bào tại điểm kiểm tra chu trình tế bào này.

Yếu tố *p53* còn có vai trò trong việc điều hòa phản ứng của tế bào trong quá trình sửa chữa ADN sai hỏng. Quá trình này được điều khiển bằng hai con đường độc lập: (i) lượng protein *p53* trong nhân tăng lên sau khi ADN bị hư hại. Điều này làm ngăn cản tế bào chuyển vào pha S cho đến khi sự hư hỏng của ADN được sửa chữa; (ii) bằng con đường khác, *p53* liên quan đến việc kích thích làm tế bào chết thông qua quá trình **chết theo chương trình**.

Cả hai con đường trên đây đều có tác dụng làm giảm nguy cơ xảy ra các sai hỏng ADN tạo ra các dạng đột biến làm tế bào phát triển thành các khối u. Nhờ những chức năng này, *p53* được mô tả là "yếu tố bảo vệ hệ gen".

Chuột chuyển gen thiếu gen *p53* hoạt động và phát triển bình thường nhưng thường có sự hình thành các khối u tự phát với tần số cao hơn rõ rệt. Trong các tế bào, mức độ khuếch đại của các gen tăng cao.

Các dạng ung thư di truyền

Một số dạng ung thư hiếm gặp và một tỷ lệ nhỏ các dạng ung thư phổ biến biểu hiện tập trung ở một số dòng họ. Những dòng họ này được coi là có xu hướng di truyền các dạng ung thư đó. Các dạng ung thư có xu hướng di truyền gặp phải có thể là một loại bệnh ung thư đặc trưng, ví dụ như ung thư vú, nhưng cũng có thể xuất hiện với nhiều bệnh ung thư khác nhau.

Xu hướng di truyền các bệnh ung thư có thể xảy ra do các đột biến ở các gen ức chế khối u trong các dòng tế bào sinh dục. Vì vậy, tất cả các tế bào sinh dưỡng của cá thể con sẽ mang một alen đột biến, và nguy cơ hình thành các tế bào khối u tăng cao lên nhiều do khả năng bất hoạt của các alen bình thường còn lại duy nhất. Ví dụ ở người có gen của u nguyên bào lưới, *Rb*, trên nhiễm sắc thể số 13; và một gen có xu hướng di truyền với bệnh ung thư vú và buồng trứng là *BRCA1* nằm trên nhiễm sắc thể số 17. Alen bình thường bị bất hoạt trong các tế bào của khối u. Các gen ung thư di truyền trong dòng họ thường tìm thấy trải qua quá trình đột biến trong tế bào sinh dưỡng trong một số trường hợp rải rác (không di truyền) của cùng một loại ung thư.

Các dạng ung thư di truyền thường biểu hiện hình thức di truyền giống di truyền trội nằm trên nhiễm sắc thể thường, do các bệnh nhân có xu hướng truyền gen gây bệnh cho một nửa số con cái, tuy vậy các đột biến trong các gen ức chế khối u thường là lặn. Điều này dường như là một nghịch lý. Nó được giải thích là do: nếu một cá thể chỉ có một bản sao của một gen ức chế khối u hoạt động trong mọi tế bào thì có nhiều khả năng, trong một số tế bào, một đột

biến thứ hai sẽ gây ra sự bất hoạt của bản sao có chức năng hoạt động bình thường đó. Những tế bào bị "đột biến kép" này có thể phát triển thành các khối u.

Các gen sửa chữa ADN cũng có thể liên quan đến hiện tượng di truyền của ung thư. Các ví dụ về hiện tượng này bao gồm bệnh sắc tố xesoderma do một đột biến lặn trên nhiễm sắc thể thường gây ra, trong đó những cá thể bị bệnh có tính mẫn cảm cao với ánh sáng mặt trời do bị thiếu các enzym endonuclease liên quan đến quá trình sửa chữa ADN. Các alen đột biến của một vài gen khác có thể dẫn đến xu hướng bị ung thư ruột kết. Hai trong số alen như vậy cũng tìm thấy có liên quan đến quá trình sửa chữa lắp ghép sai nucleotid trong quá trình sao chép ADN. Các tế bào của khối u sửa chữa ADN hư hỏng kém hiệu quả hơn so với các tế bào bình thường và tạo ra các đột biến ở tần số cao, và một số đột biến xảy ra trong các gen ức chế khối u, hoặc các gen gây ung thư sơ khai. Hiện tượng này gọi là **kiểu hình của thể đột biến.**

VI.C PHÉP THỬ ĐỘC TÍNH GEN

Phép thử độc tính gen

Đây là một lĩnh vực nghiên cứu độc học mới, tập trung nghiên cứu về các tác nhân gây độc trong môi trường thông qua việc làm hư hỏng cấu trúc ADN và vì vậy gây ra các dạng đột biến. Sự hiểu biết về những hợp chất này có ý nghĩa quan trọng do chúng gây ung thư và được gọi là **các chất gây ung thư**. Các hợp chất này gây tăng tần số đột biến trong quần thể nói chung. Việc xác định các hợp chất gây đột biến có ý nghĩa quan trọng trong nhiều lĩnh vực khác nhau, bao gồm dược phẩm, công nghệ thực phẩm (phụ gia thực phẩm), nông nghiệp, phân tích ô nhiễm môi trường và trong nhiều quá trình công nghiệp khác. Vì lý do này, việc phát triển các phép thử xác định các hợp chất nguy hiểm là một yêu cầu cần thiết. Các hệ thống phép thử trước kia dựa trên các động vật thí nghiệm cỡ nhỏ như chuột, nhưng nói chung các phép thử này tốn nhiều thời gian, chi phí cao và thường bị phản đối về khía cạnh đạo đức sinh học. Vì lý do đó, gần đây một loạt các phép thử *invitro* sử dụng vi khuẩn, hoặc tế bào động vật nuôi cấy được dùng thay thế. Ở nhiều nước, việc các hợp chất mới trước khi được cấp phép lưu thông bắt buộc phải trải qua các phép thử độc tính gen do các cơ quan hữu trách của Chính phủ thực hiện đã được ghi thành Luật.

Đến nay, có nhiều phương pháp thử khác nhau để xác định nhanh các chất gây đột biến, nhưng ở đây chỉ đề cập đến 3 phương pháp thông dụng.

Phép thử Ames

Đây là phép thử xác định nhanh các hợp chất gây đột biến. Phép thử này sử dụng vi khuẩn là loài *Salmonella typhymurium* bị đột biến ở operon histidin nên không có khả năng tổng hợp histidin, gọi là dòng dị dưỡng histidin thụ động (*his*$^-$). Các hợp chất gây đột biến được xác định nhờ khả năng chuyển dạng đột biến này trở về dạng bình thường. Phép thử này có thể thực hiện một cách đơn giản bằng việc đặt dòng vi khuẩn *his*$^-$ lên môi trường thạch chỉ chứa một lượng vết của histidin. Một hạt tinh thể hoặc một tấm giấy lọc được làm ướt bằng dung dịch chứa hợp chất

thử được đặt lên bề mặt agar. Vi khuẩn sẽ phát triển trong một thời gian ngắn cho đến khi histidin trong môi trường được sử dụng hết. Sau thời gian này, chỉ các vi khuẩn có gen *his*⁻ được chuyển hóa về dạng his^+ có khả năng tổng hợp histidin mới có khả năng phát triển và hình thành khuẩn lạc, gọi là dòng sinh dưỡng histidin (his^+). Nếu chất thử không phải chất gây đột biến thì chỉ có một vài dòng vi khuẩn chuyển hóa được tìm thấy nằm rải rác trên bề mặt thạch. Nếu chất thử là chất gây đột biến, số khuẩn lạc sẽ tăng lên và nằm thành từng nhóm xung quanh điểm được bổ sung thuốc thử. Rõ ràng phép thử này được xây dựng dựa trên nguyên tắc định lượng để xây dựng đường cong liều lượng - phản ứng đối với mỗi chất thử.

Nhằm xác định số lượng tối đa các chất gây đột biến, người ta sử dụng hai dạng đột biến his^-, một dạng do sự thay thế base nitơ duy nhất và dạng kia là đột biến dịch khung. Phương pháp này cho phép xác định các chất gây đột biến gây hậu quả khác nhau trên phân tử ADN. Nó cũng có thể làm thay đổi về mặt di truyền khả năng hấp thu chất thử của vi khuẩn, và làm giảm khả năng sửa chữa ADN hư hỏng của chúng. Điều này lại giúp việc xác định hoạt tính gây đột biến ADN dễ dàng hơn.

Đến nay, một số hợp chất được biết chỉ gây ung thư sau khi bị chuyển hóa bởi enzym trong cơ thể. Các hợp chất này được gọi là các **chất gây ung thư sơ khai**. Hoạt động của enzym trong cơ thể sẽ chuyển hóa chúng thành **chất gây ung thư cuối cùng**. Nếu các chất gây ung thư sơ khai được sử dụng trong phép thử Ames, nó sẽ cho ra kết quả âm tính, vì vậy phép thử này có thể được sử dụng để xác định các chất gây ung thư sơ khai. Gan là một mô chứa nhiều enzym hoạt động. Các dịch chiết gan chứa nhiều enzym hoạt động có thể được dùng bổ sung vào thuốc thử. Sự hoạt hóa của các chất gây ung thư sơ khai sẽ xảy ra và vì vậy số tế bào vi khuẩn his^- được chuyển hóa về dạng his^+ tăng lên.

Phép thử Ames là phép thử nhanh. Thời gian tiến hành một phép thử trong vòng 48 giờ. Đây cũng là phương pháp có chi phí thấp và dễ định lượng. Bằng phương pháp này người ta đã xác định được nhiều chất gây đột biến khác nhau, bao gồm một số loại thuốc nhuộm tóc, chất cứu hỏa và phụ gia tạo màu ở thực phẩm.

Phép thử gây đột biến dòng tế bào

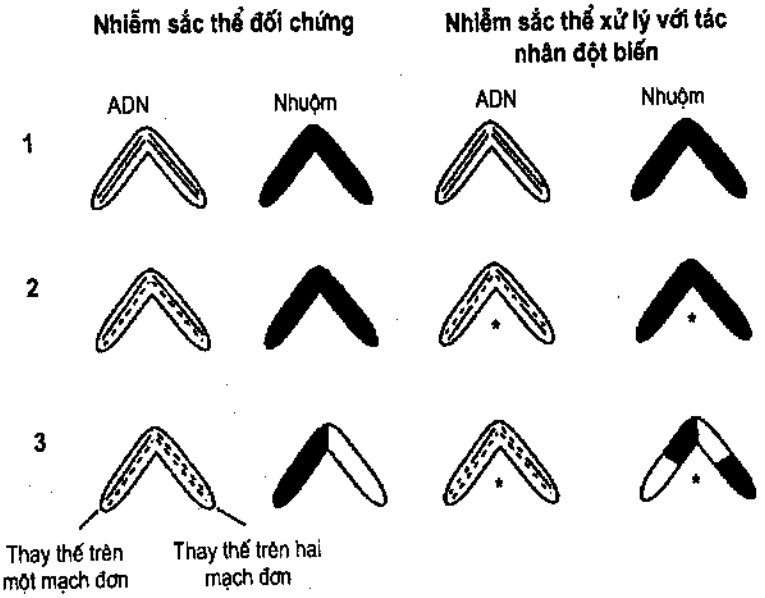
Một nhược điểm của phép thử Ames là tế bào thử đột biến là tế bào vi khuẩn, không phải tế bào động vật có vú. Vì lý do này, một số phép thử khác được phát triển dựa trên việc sử dụng các dòng tế bào động vật có vú được nuôi cấy. Trong số các dòng này, quan trọng nhất là dòng dị hợp tử về gen chuyển hóa thymidin L5178Y. Đây là một dòng tế bào thu được từ lympho bào của chuột. Nó mang hai alen chức năng ở locut gen mã hóa enzym chuyển hóa thymidin *tk*. Đột biến gây thiếu năng về enzym này (tk^-) xuất hiện trong tự nhiên với tần số rất thấp để có thể sử dụng trong một hệ thống thử, nhưng dạng dị hợp tử $tk^{+/-}$ được phân lập có tần số đột biến về dạng tk^- có thể định lượng được. Các dạng đột biến có thể được xác định bằng việc chọn lọc các tế bào được xử lý với các chất đồng đẳng gây độc của thymidin. Chỉ các tế bào tk^- sống sót và hình thành khuẩn lạc. Tính chất này được sử dụng giống như nguyên tắc trong phép thử Ames, cũng như có thể bổ

sung thêm các dịch chiết từ mô gan. Tuy vậy, do sự sinh trưởng chậm của các tế bào động vật trong quá trình nuôi cấy, phép thử kéo dài đến 3 tuần mới thu nhận được kết quả.

Nhiều phép thử khác được xây dựng dựa trên các điểm dừng khác nhau của tế bào động vật trong điều kiện *in-vivo* hoặc *in-vitro*.

Phép thử di truyền học tế bào

Các hợp chất gây hư hỏng nhiễm sắc thể gọi là **các chất gây đột biến nhiễm sắc thể** (clastogen). Các hợp chất này cũng có thể phát hiện được nhờ hệ thống các phép thử. Các phép thử này được tiến hành trên các dòng tế bào của động vật thí nghiệm, thậm chí ở thực vật. Các phép thử này bao gồm các dạng khác thường của nhiễm sắc thể có thể đo được như các dạng đứt gãy nhiễm sắc thể, chuyển đoạn nhiễm sắc thể, nhiễm sắc thể vòng, lặp đoạn nhiễm sắc thể và đảo đoạn nhiễm sắc thể. Tần số biến dạng nhiễm sắc thể thường rất thấp và việc xác định được tần số biến dạng cần phải có các kỹ thuật viên có trình độ cao. Một phương pháp khác được dùng để liệt kê các biến dạng nhiễm sắc thể lớn là phương pháp **trao đổi nhiễm sắc tử chị em SCE** (*s*ister *c*hromatid *e*xchanges). SCE dựa trên sự trao đổi vật chất giữa hai nhiễm sắc tử chị em của cùng một nhiễm sắc thể, và là quá trình xảy ra tự nhiên với tần số thấp trong mọi tế bào. Quá trình này xảy ra trong cả nguyên phân và giảm phân. Khi tế bào được xử lý với chất gây đột biến nhiễm sắc thể, tần số SCE tăng lên rõ rệt và tạo ra các biến dạng nhiễm sắc thể. Nhờ vậy, kết quả phân tích có thể thu được từ phép phân tích một số lượng nhỏ tế bào và có thể được thực hiện bởi các kỹ thuật viên thông thường. SCE có thể được xác định bằng một số quy trình kỹ thuật khác nhau, nhưng tất cả đều dựa trên bản chất bán bảo toàn của quá trình sao chép ADN. Ở đây chỉ đề cập đến phương pháp sử dụng chất nhuộm Giêmsa. Bản chất của kỹ thuật này là nhuộm màu các nhiễm sắc tử chị em với thuốc nhuộm khác nhau, nhờ vậy dễ dàng nhận ra sự trao đổi chéo giữa các nhiễm sắc tử chị em. Các tế bào được nuôi cấy trong hai vòng sao chép trên môi trường có nồng độ thấp không gây độc của hợp chất đồng đẳng của thymidin là bromodeoxyuridin (BrdU). Chất này sẽ liên kết vào sợi ADN mới được tổng hợp thay thế cho vị trí thymidin và làm thay đổi thuộc tính nhuộm màu của nhiễm sắc tử. Sau một chu kỳ sao chép, cả hai nhiễm sắc tử trong phân tử ADN sợi kép đều mang một sợi chứa BrdU còn sợi kia chứa thymidin. Cả hai nhiễm sắc tử lúc này gọi là các nhiễm sắc tử được **thay thế trên một mạch đơn**. Việc nhuộm các nhiễm sắc tử được thực hiện giống như việc nhuộm nhiễm sắc thể từ các tế bào không được xử lý với BrdU. Sau hai chu kỳ sao chép, mỗi nhiễm sắc thể chứa hai nhiễm sắc tử, một trong số chúng được thay thế T bằng BrdU theo kiểu thay thế trên một mạch đơn, nhưng phân tử kia chứa cả hai mạch bị thay thế T bằng BrdU, gọi là nhiễm sắc tử được **thay thế trên hai mạch đơn**. Khi nhiễm sắc tử được thay thế trên hai mạch đơn này được nhuộm Giêmsa, nhiễm sắc tử bị thay thế trên một mạch đơn bắt màu đậm, trong khi nhiễm sắc tử bị thay thế trên hai mạch đơn bắt màu rất nhạt. Nhờ vậy, lúc đó có thể dễ dàng phát hiện được các đoạn nhiễm sắc thể được trao đổi giữa hai nhiễm sắc tử (*Hình 6.7*). Quá trình nhuộm tương đối phức tạp, liên quan đến việc chụp chất phát quang dưới nguồn sáng UV và sau đó nhuộm bằng thuốc nhuộm Giêmsa.



Hình 6.7. Hình ảnh mô tả hai bộ nhiễm sắc thể (NST). Bên phải là các NST được xử lý với tác nhân đột biến để tạo ra sự trao đổi nhiễm sắc tử chị em (SCE), còn bên trái là các NST đối chứng không được xử lý. Mỗi NST được minh họa bằng trạng thái ADN và kiểu hình bắt màu thuốc nhuộm. Đường liền nét biểu diễn mạch ADN đơn mang thymin, đường đứt quãng biểu diễn mạch ADN đơn mang bromodeoxyuridin. 1. Trước khi được đánh dấu với bromodeoxyuridin, cả hai nhiễm sắc tử đều mang ADN chỉ có thymin, và nhuộm màu sẫm. 2. Sau một chu trình tế bào với bromodeoxyuridin, trong mỗi NST phân tử ADN mang một mạch chứa bromodeoxyuridin còn mạch kia chứa thymin; tất cả các nhiễm sắc tử bắt màu thuốc nhuộm sẫm. Một SCE xảy ra trong NST được xử lý với tác nhân đột biến trong chu trình tế bào này tại điểm tương ứng với dấu *. 3. Sau chu trình tế bào thứ hai với sự có mặt của bromodeoxyuridin, NST đối chứng chứa các phân tử ADN được đánh dấu khác nhau trên mỗi nhiễm sắc tử. Một nhiễm sắc tử bị thay thế trên một mạch đơn với bromodeoxyuridin, và bắt màu thuốc nhuộm đậm; nhiễm sắc tử kia bị thay thế trên cả hai mạch đơn và bắt màu thuốc nhuộm nhạt. Hiện tượng tương tự cũng xảy ra với NST được xử lý đột biến và nhờ vậy lúc này có thể quan sát được sự trao đổi giữa các nhiễm sắc tử chị em (SCE).

VI.D LIỆU PHÁP GEN

Tổng quan

Sự tiến bộ vượt bậc về các kỹ thuật di truyền phân tử trong 20 năm qua đã giúp chúng ta có thể phân lập, tách dòng và giải mã trình tự các gen từ tất cả các loài sinh vật, kể cả loài người. Bên cạnh những khả năng này, việc có thể lập bản đồ hệ gen người nói chung và các gen liên quan đến các bệnh di truyền nói riêng cho phép chúng ta có thể xác định được các nguyên nhân di truyền một cách rõ ràng hơn đối với các dạng bệnh lý ở nhiều cá thể khác nhau. Kết quả đó dẫn đến một câu hỏi là liệu có thể sử dụng các biện pháp di truyền để điều trị cho các bệnh nhân bị bệnh di truyền, hoặc ít nhất là cải thiện điều kiện sức khỏe của họ hay không. Điều này đòi hỏi phải phân lập được gen có liên quan và chuyển vào tế bào đích của cá thể mang bệnh, phương pháp như vậy được gọi là **liệu pháp gen**. Trong trường hợp các hội chứng di truyền, có thể có một số phương thức để lựa chọn: một bản sao bình thường của gen có thể được biến nạp hoặc vào dòng tế bào sinh dục, hoặc vào dòng tế bào sinh dưỡng của cơ thể bị bệnh. Việc chuyển gen

vào tế bào sinh dục sẽ tránh được việc truyền các alen hư hại sang thế hệ sau, trong khi việc chuyển gen vào tế bào sinh dưỡng có thể cải thiện được hiệu quả của alen ngay trong cơ thể bị bệnh. Các nghiên cứu hiện nay chủ yếu liên quan đến việc xử lý trên các tế bào sinh dưỡng. Trong phần lớn các trường hợp, một bản sao bình thường của gen được chuyển vào các tế bào có hai alen không hoạt động, nhờ vậy có thể sửa chữa ngay trạng thái sai hỏng di truyền đang tồn tại. Tuy vậy, trong liệu pháp gen được đề xuất cho việc điều trị bệnh ung thư, có thể chuyển các gen vào các tế bào khối u nhằm làm chết các tế bào này. Nếu như các vấn đề liên quan đến việc tách dòng các gen tương ứng và việc điều hòa sự biểu hiện của chúng được giải quyết, khó khăn chính trong việc ứng dụng thành công liệu pháp gen là việc tìm ra vector phù hợp để chuyển gen vào tế bào đích một cách hiệu quả mà không gây các tác động phụ. Một nghịch lý là vector có tiềm năng nhất sử dụng trong liệu pháp gen lại xuất phát từ các dạng virus gây bệnh ở người, sở dĩ như vậy bởi vì chúng đã phát triển một hệ thống biến nạp gen hiệu quả vào hệ gen của tế bào người.

Liệu pháp gen có thể thực hiện được bằng việc biến nạp gen vào bệnh nhân và hy vọng chúng sẽ tìm thấy tế bào đích - đây là **liệu pháp *in vivo***. Rõ ràng đây là một quá trình không hiệu quả. Trong một số trường hợp, người ta có thể phân lập được các tế bào từ bệnh nhân, thao tác chúng trong điều kiện nuôi cấy tế bào, sau đó đưa chúng trở lại cơ thể người bệnh. Quá trình này gọi là **liệu pháp *ex vivo***.

Một điều quan trọng có thể dễ dàng nhận thấy là có nhiều loại **tế bào đích** khác nhau. Trong các trường hợp nhất định, gen phải được biến nạp vào đúng loại tế bào đặc thù, ví dụ việc thay thế các gen globin trong điều trị các bệnh lý di truyền liên quan đến hemoglobin miễn dịch phải được chuyển vào các tế bào hồng cầu. Trong một số trường hợp, sự biểu hiện của một loại protein đặc thù nào đó trong cơ thể người bệnh phải được tăng cường. Trong những trường hợp này, các loại tế bào cho phép giải phóng sản phẩm phiên mã của gen liệu pháp vào hệ thống mạch máu được lựa chọn làm tế bào đích.

Vector chuyển gen

Trước khi virus có thể được sử dụng làm vector chuyển gen liệu pháp, các gen cần thiết cho hoạt động của virus (gen gây bệnh) phải được loại bỏ. Điều này giúp tránh virus sinh sôi và phát tán trong cơ thể, đồng thời hạn chế phản ứng miễn dịch ở cơ thể người bệnh được điều trị. Virus được cải biến theo cách này được gọi là các loại **virus bị rút lõi**. Các gen liệu pháp cần được biến nạp sẽ được cài vào phần hệ gen còn lại của virus cùng với các trình tự ADN giúp khởi động và điều hòa sự phiên mã của gen, và phân tử ADN tái tổ hợp này được đóng gói vào vỏ virus thích hợp, nhờ vậy nó có thể lây nhiễm vào tế bào người. Đến nay, có nhiều loại vector virus khác nhau đã được phát triển cho mục đích liệu pháp gen.

Retrovirus là dạng virus dễ thích nghi với việc mang một đoạn cài gen của người. Các loại virus này còn có các ưu điểm là có thể lây nhiễm phần lớn (nếu không muốn nói là gần như tất cả) các loại tế bào và khả năng kết hợp hiệu quả hệ gen của chúng với hệ gen của tế bào chủ sau khi lây nhiễm. Nhược điểm tồn tại của chúng là chỉ lây nhiễm các tế bào đang phân chia mạnh. Vì nguyên nhân này, việc sử dụng dạng virus này chủ yếu giới hạn ở liệu pháp *ex vivo* tiến hành trên các tế bào được phân lập và nuôi cấy. Một vấn đề khác là các đoạn gen được biến nạp vào vector này chỉ được biểu hiện trong một khoảng thời gian ngắn (thường

ít hơn 2 tuần), gọi là biểu hiện tạm thời. Một phân nhóm của retrovirus, gọi là **lentivirus**, trong đó có virus gây hội chứng suy giảm miễn dịch mắc phải ở người HIV, đã được nghiên cứu phát triển nhằm sử dụng trong liệu pháp gen. Dạng virus này có thể lây nhiễm cả các tế bào đang ngừng phân bào, vì vậy có thể áp dụng trong các quy trình *in vivo*.

Một nhóm virus có hệ gen là ADN cũng được sử dụng trong các kỹ thuật liệu pháp gen là nhóm **adenovirus**. Một số dạng adenovirus liên quan đến các bệnh viêm nhiễm đường hô hấp trên ở người. Hệ gen của chúng có thể mang một đoạn cài ADN lớn của người, nhưng chúng cũng có những nhược điểm trong đó bao gồm việc hệ gen của chúng không kết hợp được vào hệ gen tế bào người. Điều này có nghĩa là chúng không thể tạo ra hiệu quả điều trị vĩnh viễn và triệt để trong các tế bào chúng lây nhiễm. Chúng cũng có thể kích ứng hệ thống miễn dịch và dẫn đến việc làm chết tế bào bị lây nhiễm. Nhưng cũng nhờ vậy, chúng có thể được dùng vào việc tiêu diệt các khối u. Các loại virus ADN khác đang được thử nghiệm làm vector dùng trong liệu pháp gen bao gồm các loại vector liên quan đến adenovirus và virus *Herpes simplex*, một loại virus vốn có một hệ gen lớn và nhờ vậy có thể mang các đoạn cài có kích thước lớn.

Tất cả các vector virus điều có các hạn chế, hoặc là do hiệu quả biến nạp gen thấp, hoặc là do khả năng kích ứng hệ thống miễn dịch, hoặc là do chúng có liên hệ gần gũi với các dạng virus có khả năng gây khối u. Vì nguyên nhân này, nhiều nỗ lực nghiên cứu đã được tiến hành nhằm phát triển các hệ thống biến nạp gen liệu pháp không sử dụng vector virus. Trong số này, phương tiện chuyển gen triển vọng nhất là **liposom** (còn gọi là thể mỡ) có thể mang các đoạn ADN trần vượt qua màng tế bào. Phân tử phospholipid gắn vào phân tử ADN và vì vậy nó có thành phần giống với thành phần của màng tế bào và hai cấu trúc có thể dung hợp vào nhau, nhờ vậy ADN được chuyển vào tế bào chất của tế bào đích. Liposom có một số ưu điểm: chúng không kích ứng hệ thống miễn dịch; chúng có thể hoạt động trong cả hai điều kiện *ex vivo* và *in vivo*, và chúng có thể mang phân tử ADN ở mọi kích thước khác nhau. Tuy vậy, phân tử ADN mà nó vận chuyển có tần số kết hợp với hệ gen tế bào nhận thấp. Đến nay, đã có hàng trăm nghiên cứu thử nghiệm lâm sàng được tiến hành nhằm sử dụng liệu pháp gen trong điều trị các bệnh ở người, và đồng thời các hệ thống mô hình động vật thí nghiệm cũng được dùng để xác định hiệu quả của các phương pháp mới. Dưới đây là một số ví dụ về sử dụng liệu pháp gen.

Thiểu năng enzym adenin deaminase

Đây là một bệnh di truyền hiếm gặp do gen lặn nằm trên nhiễm sắc thể thường quy định, trong đó người bệnh mang hai alen đột biến thuộc locut gen mã hóa enzym chuyển hóa adenin ADA (adenin deaminase). Do các hậu quả sai hỏng trong các enzym chuyển hóa nucleotid, việc thiếu hụt về enzym này có các ảnh hưởng khác nhau trên các loại mô khác nhau. Việc thiếu hụt ADA có hậu quả nghiêm trọng nhất trên các tế bào thuộc hệ thống miễn dịch, dẫn đến sự phát triển của hội chứng thiểu năng miễn dịch phối hợp cấp tính. Gen mã hóa ADA được tách dòng phân tử và những nỗ lực điều trị bệnh thiểu năng ADA đầu tiên được thực hiện vào năm 1990. Gen này được cài vào vector retrovirus, tế bào lympho T được tách ra từ người bệnh, được kích thích phát triển trong điều kiện nuôi cấy tế bào và được biến nạp với gen ADA bình thường. Các tế bào mang gen biểu hiện chức năng bình thường, sau đó được nhân lên bằng quá trình nuôi cấy rồi được chuyển trở lại cơ thể người bệnh. Quy trình này tỏ ra thành công ở một số trường hợp. Nhưng một quy

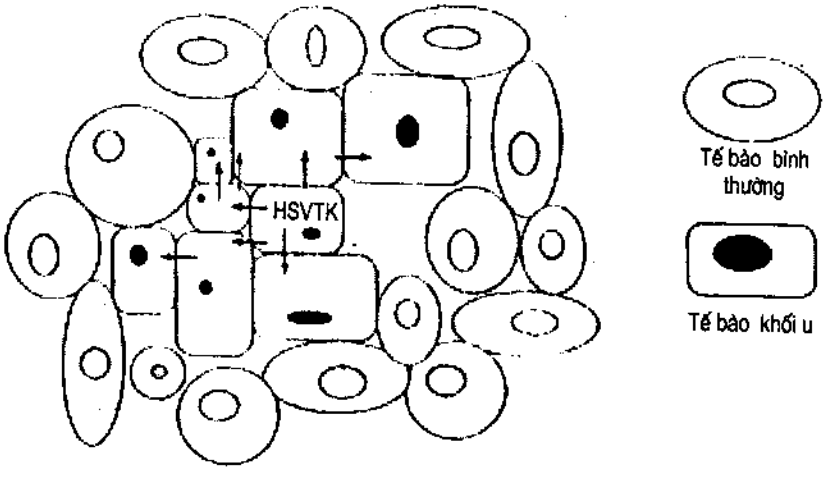
trình điều trị đơn lẻ không chữa được triệt để căn bệnh này do tế bào lympho T là tế bào đã biệt hóa trưởng thành, chúng sẽ bị thay thế bởi tế bào mới được sản sinh từ tủy xương, vì vậy quy trình điều trị trên đây phải được lặp đi lặp lại nhiều lần để có thể duy trì lượng ADA cần thiết trong cơ thể người bệnh. Nếu như tế bào gốc sản sinh tế bào lympho T ở tủy xương được biến nạp gen hiệu quả thì quy trình điều trị bằng liệu pháp gen có thể chỉ cần thực hiện một lần duy nhất.

Bệnh xơ nang

Bệnh xơ nang, cũng giống như ADA, có nguyên nhân bắt nguồn từ một đột biến lặn đơn gen nằm trên nhiễm sắc thể thường, tuy vậy tần số đột biến của gen này cao hơn nhiều, đặc biệt trong các quần thể người ở vùng Cápca. Sự sai hỏng xảy ra trong gen có trách nhiệm vận chuyển ion Cl^- qua màng tế bào. Việc thiếu chức năng này gây hậu quả lớn lên các tế bào màng phổi và ruột. Liệu pháp gen đầu tiên được thử nghiệm cũng giống như trường hợp điều trị ADA trong đó gen được tách dòng và biến nạp vào một vector. Do bản chất của tế bào đích, liệu pháp gen *ex vivo* không thực hiện được. Vì vậy, người ta đã sử dụng adenovirus làm vector biến nạp gen. Virus được cải tiến di truyền mang một gen biểu hiện chức năng bình thường được chuyển vào cơ thể người bệnh bằng biện pháp "phun" vào đường hô hấp. Một số tế bào có thể lây nhiễm vào các tế bào của đường hô hấp. Vấn đề tồn tại trong kỹ thuật này là việc lây nhiễm các tế bào thông qua màng nhày của mô gặp khó khăn, ngoài ra adenovirus không kết hợp được hệ gen của nó vào hệ gen của tế bào chủ và phản ứng miễn dịch của tế bào chủ chống lại virus lây nhiễm có thể diễn ra. Một biện pháp thay thế là người ta sử dụng thể mỡ liposom để chuyển gen. Kỹ thuật này cho phép đưa gen vào một số ít tế bào hơn, nhưng hiệu quả điều trị kéo dài hơn.

Bệnh ung thư

Phần lớn các thử nghiệm lâm sàng sử dụng liệu pháp gen được áp dụng trên các tế bào khối u. Nếu các khối u xuất hiện do kết quả của đột biến gen (xem mục VI.B), thì về lý thuyết việc thay thế các gen ức chế khối u bị hỏng có thể ngăn cản được quá trình phát triển của các tế bào khối u. Một phương pháp khác là có thể sử dụng các gen ngoại lai hoặc để giết tế bào khối u hoặc biến các tế bào khối u trở nên mẫn cảm hơn với các cơ chế bảo vệ cơ thể bình thường.



Hình 6.8. **Sự tiêu diệt các tế bào khối u bởi gen mã hóa thymidin kinase của virus *Herpes simplex* - HSVTK (*Herpes simplex* virus thymidin kinase). Một tế bào khối u được biến nạp với gen HSVTK có thể xuất tiết các dẫn xuất phosphoryl hóa độc của gancyclovir tới các tế bào khối u khác. Các tế bào bình thường dường ít bị ảnh hưởng hơn do chúng có hoạt động tổng hợp ADN thấp**

Một phương pháp có nhiều tiềm năng khác là có thể giết các tế bào khối u một cách đặc hiệu bằng việc chuyển gen mã hoá enzym chuyển hóa thymidin (thymidin kinase) của virus *Herpes simplex*. Enzym thymidin kinase của virus khác đáng kể so với enzym tương ứng của người và có khả năng phosphoryl hóa nhiều loại cơ chất hơn. Một loại cơ chất không chuyển hóa được bởi enzym thymidin kinase của người là gancyclovir. Hợp chất này không độc với tế bào người, nhưng dẫn xuất bị phosphoryl hóa của nó rất độc. Việc chuyển gen thymidin kinase của Herpes vào tế bào khối u làm cho chúng trở nên mẫn cảm một cách đặc hiệu hơn với việc xử lý gancyclovir. Phương pháp này có hai ưu điểm, đó là: (i) enzym thymidin kinase chỉ hoạt động trong các tế bào đang sao chép ADN và vì vậy các tế bào ngoài khối u ít bị ảnh hưởng; và (ii) các phân tử gancyclovir đã bị phosphoryl hóa, vốn chỉ được tạo ra trong các tế bào khối u đã tiếp nhận gen biến nạp, được "xuất" sang các tế bào lân cận của khối u qua các kênh liên bào. Kết quả là tỷ lệ các tế bào khối u bị gancyclovir giết chết tăng lên nhiều. Kỹ thuật này đã được sử dụng trong điều trị các khối u buồng trứng và não. Sơ đồ minh họa được trình bày trên *Hình 6.8*.

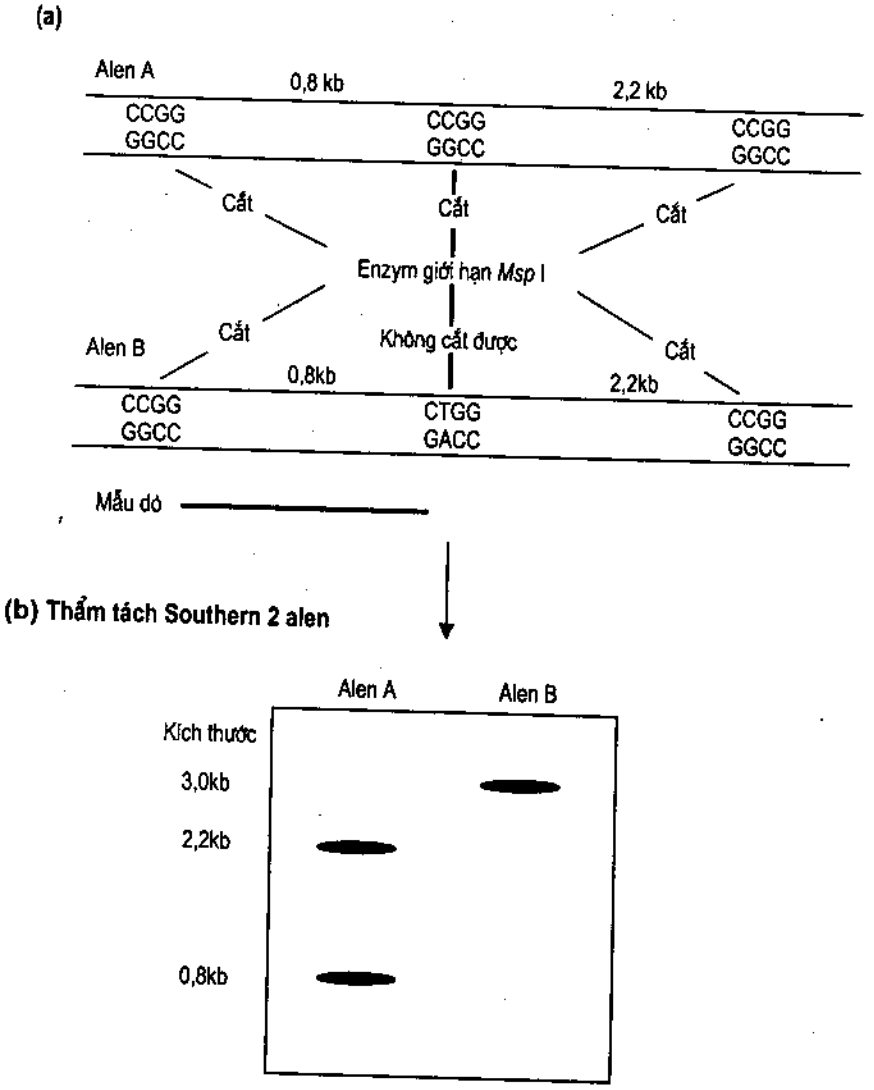
VI.E DỰ ÁN HỆ GEN NGƯỜI

Cơ sở khoa học

Vào khoảng những năm 1980-1990, toàn bộ trình tự ADN của hệ gen virus đầu tiên đã được xác định. Thành tựu này chỉ ra khả năng có thể giải mã toàn bộ hệ gen của một số sinh vật có kích thước lớn hơn, trong đó có con người. Ý tưởng này lúc đầu gây nhiều tranh cãi. Thậm chí, ngay trong nhóm những người ủng hộ việc giải mã hệ gen người cũng có những bất đồng trong việc nên hay không nên giải mã toàn bộ hệ gen (thay vào đó, chỉ giải mã những vùng gen mã hóa). Năm 1991, đánh dấu một bước khởi đầu quan trọng cho dự án nhằm giải mã toàn bộ hệ gen người dự kiến kết thúc vào năm 2005. Trách nhiệm điều phối dự án này được giao cho Cơ quan nghiên cứu hệ gen người HUGO (*H*uman *G*enome *O*rganisation). Song song với mục tiêu của dự án là giải mã hệ gen người, một số dự án giải mã hệ gen của một số loài sinh vật khác cũng đã được tiến hành. Trong số đó, một số dự án giải mã hệ gen vi khuẩn và nấm men đã hoàn thành. Trước khi công nghệ giải mã trình tự có thể phát huy tác dụng, người ta phải lập bản đồ di truyền toàn bộ hệ gen người. Việc này được thực hiện bởi hai quá trình khác biệt nhưng phối hợp với nhau gọi là kỹ thuật lập bản đồ di truyền và kỹ thuật lập bản đồ vật lý. Việc tạo ra các bản đồ khung là mục tiêu đầu tiên của dự án giải mã hệ gen người. Tốc độ thực hiện việc lập bản đồ gen sẽ phụ thuộc vào việc tìm ra các dấu chuẩn di truyền; một dấu chuẩn có thể là một đoạn ADN, một phân tử protein hoặc một kiểu hình nhất định. Nếu một dấu chuẩn có các biến thái khác nhau trong quần thể được gọi là tính đa hình. Lý tưởng nhất là các dấu chuẩn này phân bố tương đối đồng đều khắp hệ gen.

Bản đồ di truyền thô đầu tiên của hệ gen người dựa vào các kiểu hình có thể dễ nhận biết, chẳng hạn như nhóm máu và các hội chứng di truyền như Nail-Patella liên kết với nhau trên nhiễm sắc thể số 9. Trước khi các dấu chuẩn ADN được sử dụng rộng rãi, các dấu chuẩn phân tử đầu tiên dựa trên sự xác định tính đa hình của các enzym (izozym) được phát hiện bằng điện di, nhưng số lượng và sự phân bố của các dấu chuẩn di truyền này nhìn chung không đủ

để xây dựng một bản đồ khung để từ đó có thể thực hiện dự án giải mã trình tự. Tiến bộ đáng kể đầu tiên trong quá trình này phải kể đến việc xây dựng bản đồ Donis - Keller vào năm 1987. Bản đồ này sử dụng 180 locut dấu chuẩn (locut là một vị trí bất kỳ trên nhiễm sắc thể) được xác định bởi các đa hình độ dài đoạn phân cắt giới hạn (RFLP). Các dấu chuẩn RFLP được xác định như một phương pháp thay thế cho thẩm tách Southern khi một đoạn ADN được tách dòng đặc hiệu được sử dụng để dò đoạn ADN hệ gen (*Hình 6.9*): các kiểu băng khác nhau sẽ được tạo ra nếu đột biến xuất hiện trong trình tự của phân tử ADN được nhận biết bởi enzym giới hạn. Vì vậy, mọi trình tự ADN được tách dòng đều có thể sử dụng để phát hiện ra các dấu chuẩn RFLP thậm chí ngay khi không biết gì về trình tự của nó, hoặc vị trí của nó trên bản đồ di truyền (một trình tự ẩn danh). Nếu một đoạn trình tự phát hiện được kiểu hình đa hình thì nó có thể được dùng để lập bản đồ gen.

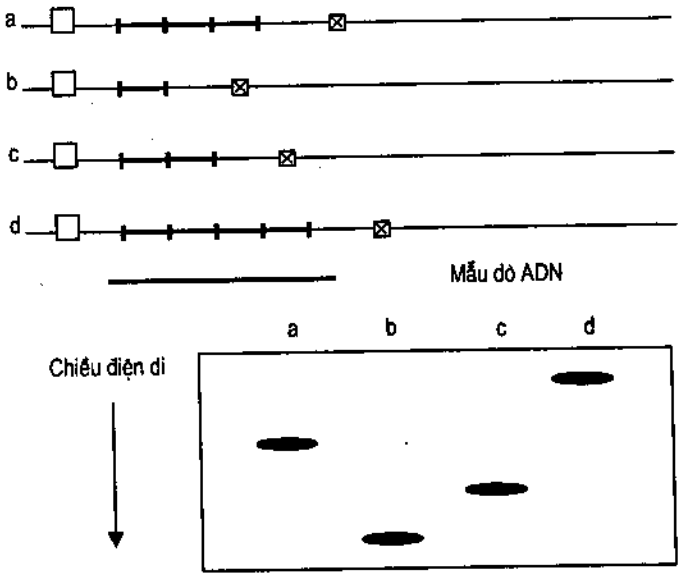


Hình 6.9. **Phát hiện chỉ thị đa hình độ dài đoạn phân cắt giới hạn (RFLP = *r*estriction *l*ength *f*ragment *p*olymorphism) ở hai alen của một locut tại vị trí cắt của enzym giới hạn *Msp*I. (a) trong hai alen thuộc hệ gen người, alen A được cắt bởi enzym giới hạn *Msp* I thành hai phân đoạn có kích thước tương ứng là 0,8 và 2,2kb. Một đột biến xảy ra ở alen B làm mất một vị trí cắt giới hạn của *Msp* I và chỉ cho ra phân đoạn có kích thước 3,0kb. (b) Các phân đoạn cắt giới hạn được phát hiện bằng phương pháp thẩm tách Southern.**

Bảng 6.1. **Trình tự của một số locut tiểu vệ tinh VNTR**

Locut	Nhiễm sắc thể	Trình tự lặp lại cơ bản[a]
YNZ2	1	GAGGCTCATGGGGCACA
THH59	17	CTGGGGAGCCTGGGGACTTTTCCACACC
α- globin	18	AACAGGGACACGGGGGG

[a]Những trình tự này xuất hiện với số lần lặp lại khác nhau trong các alen khác nhau của cùng locut



Hình 6.10. **Phát hiện chỉ thị đoạn trình tự lặp lại với số lượng biến động (VNTR). a, b, c và d biểu diễn bốn bản sao của một nhiễm sắc thể trong đó vùng được vẽ bằng nét — là vùng có mức độ lặp lại với số lượng khác nhau. □ và ⊠ là các vị trí cắt giới hạn tại vùng biên của đoạn trình tự lặp lại kế tiếp. Khi ADN được cắt ở vị trí này, và sử dụng mẫu dò để phân tích theo phương pháp thẩm tách Southern thì sẽ thu được các băng điện di có kích thước khác nhau tương ứng với số lần lặp lại của trình tự cơ bản.**

Các dấu chuẩn RFLP thường hiếm, và tính đa hình thường không đủ lớn để có thể sử dụng lập bản đồ chi tiết. Một cải tiến đáng kể so với phương pháp làm dấu chuẩn RFLP là việc phát hiện ra các trình tự lặp lại liên tiếp có số lượng biến động VNTR (*v*ariable *n*umber *t*andem *r*epeat sequences) trong hệ gen của sinh vật nhân chuẩn. Đây là những trình tự lặp lại xuất hiện trong hệ gen như những nhóm trình tự liên tiếp. Số lần lặp lại của mỗi nhóm thường khác nhau giữa các bản sao của cùng một nhiễm sắc thể, vì vậy chúng xuất hiện với mức độ đa hình cao. Các dấu chuẩn VNTR được chia làm hai nhóm: (i) nhóm lặp lại của các trình tự lớn, thường bao gồm từ 10 - 25 base nitơ, được gọi là dấu chuẩn tiểu vệ tinh - minisatellite (*Bảng 6.1*). Mức độ biến động của nhóm này được phát hiện bởi các băng có kích thước khác nhau trên bản gel được thẩm tách theo phương pháp Southern như được minh họa trên *Hình 6.10*. Các dấu chuẩn tiểu vệ tinh làm tăng đáng kể số lượng các dấu chuẩn đa hình, nhưng sự phân bố của chúng không ngẫu nhiên; chúng thường được gặp nhiều hơn ở vùng tâm mút của

các nhiễm sắc thể ở người; (ii) nhóm lặp lại với các đơn vị lặp lại gồm từ 1 - 4 base nitơ được gọi là các dấu chuẩn vi vệ tinh (microsatellite). Các đoạn lặp lại gồm 2 nucleotid (CA) là dạng dấu chuẩn ADN phổ biến nhất biểu hiện bằng mức độ đa hình cao và phân bố khắp hệ gen. Các dấu chuẩn vi vệ tinh chỉ gồm 1, 3 hoặc 4 nucleotid cũng được sử dụng nhưng xuất hiện với tần số thấp hơn. Các dấu chuẩn vi vệ tinh được phát hiện bằng phản ứng chuỗi trùng hợp PCR, thay cho phương pháp thẩm tách Southern: việc này làm giảm lượng ADN cần dùng cho quá trình phân tích và thực hiện quá trình phân tích nhanh hơn. Việc phân tích bằng PCR đòi hỏi phải có một vùng ADN đặc trưng duy nhất ở hai đầu (vùng biên) của đoạn trình tự lặp lại cần giải mã trình tự để có thể cung cấp được các đoạn mồi đặc hiệu. Số lần lặp lại của các trình tự sẽ làm thay đổi kích thước sản phẩm của phản ứng PCR. Ví dụ về một trình tự vi vệ tinh điển hình được trình bày trên *Hình 6.11*.

CCACCTCCCCTCAACCATGGTTCCCACTCACATAACTCCCCCCACACTCTGTACCCTCCATGGC
CCCCAGCACACACATGCACATGCACACCTACACGGGCACATTTGCACACATT**CACACACACAC**
ACACACACACACACACACACACACACACAGAGCCATGGCCTCTGCTGGCAG

Hình 6.11. **Trình tự gạch chân là vị trí hai đoạn mồi được dùng để khuếch đại đoạn trình tự vi vệ tinh bằng kỹ thuật PCR. Vùng trình tự CA lặp lại giữa hai đoạn mồi được in đậm. Số lượng cặp nucleotid [CA] lặp lại biến động giữa các alen khác nhau, vì vậy sản phẩm PCR của các alen khác nhau có thể phân biệt được bằng phương pháp điện di.**

Bản đồ di truyền

Bản đồ di truyền được thiết lập dựa trên tần số tái tổ hợp và được xây dựng theo phương pháp đã được mô tả ở mục C4. Tuy vậy, do tính đa hình chỉ thị ADN là đồng trội, nên kiểu gen của một cá thể có thể xác định được trực tiếp, không giống như kiểu hình. Mặc dù bản đồ di truyền rất hữu ích vì nó chỉ ra vị trí của các gen trên một nhiễm sắc thể, nhưng nó không cho biết khoảng cách chính xác giữa các locut. Sở dĩ như vậy bởi vì tần số trao đổi chéo không phân bố đồng đều tuyệt đối trên nhiễm sắc thể: tần số trao đổi chéo cao ở vùng gần tâm mút và giảm đi ở vùng gần tâm động và trên nhiễm sắc thể có những vùng có tần số trao đổi chéo cao hơn so với những vùng khác. Mặc dù vậy, bản đồ di truyền cho thấy vị trí chính xác của các gen trên nhiễm sắc thể. Một ví dụ về bản đồ di truyền là công trình do nhóm các nhà khoa học Pháp (Genethon) tiến hành với các chỉ thị vi vệ tinh ở dòng họ CEPH (một dòng họ lớn được dùng làm chuẩn để xây dựng bản đồ hệ gen người) với tổng số 5236 locut vi vệ tinh được xác định vị trí. Khoảng cách trung bình giữa các dấu chuẩn này là 1,6cM, trong đó khoảng cách lớn nhất là 11cM. Những dấu chuẩn này được sử dụng làm bản đồ khung để lập bản đồ chính xác vị trí các gen gây bệnh và là cơ sở cho việc giải mã trình tự hệ gen.

Bản đồ vật lý

Bản đồ vật lý cũng phụ thuộc vào khả năng trao đổi chéo giữa các locut nhưng khoảng cách vật lý của chúng được xác định trên phân tử ADN. Thực chất, bản đồ vật lý được thiết lập bằng việc chia nhỏ hệ gen người thành các phân đoạn ngày càng nhỏ hơn và xác định xem những locut nào cùng liên kết trong các phân đoạn đó. Một dạng bản đồ vật lý đơn giản đã được xác lập khi người ta sử dụng các tế bào lai soma giữa tế bào người và tế bào chuột để xác định các gen trên các nhiễm sắc thể đặc trưng của người.

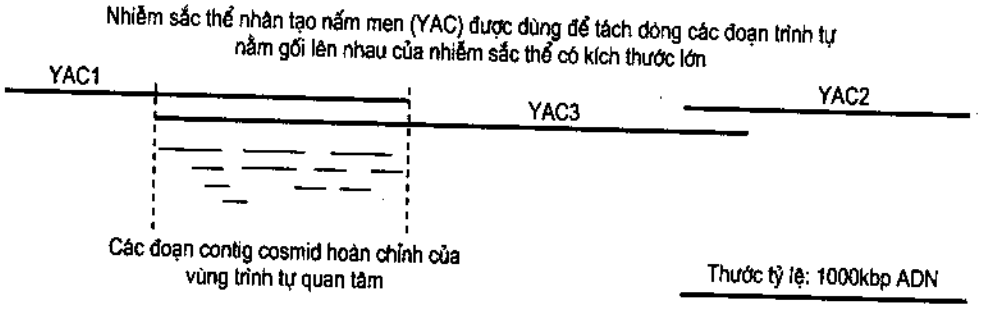
Sự phát triển của kỹ thuật **lai chiếu xạ** đã tạo điều kiện cho việc xây dựng các bản đồ vật lý cấu trúc. Trong phương pháp này, các tế bào có nguồn gốc người được chiếu xạ tia X trước khi dung hợp nhằm làm gẫy các nhiễm sắc thể. Các tế bào chết sau đó được phục hồi nhờ việc dung hợp với một dòng tế bào có nguồn gốc từ chuột. Trong tế bào lai, các phân đoạn nhiễm sắc thể người sẽ kết hợp vào nhiễm sắc thể chuột, còn các phân đoạn không liên kết sẽ bị mất đi. Do quá trình kết hợp nhiễm sắc thể xảy ra ngẫu nhiên, nên mỗi dòng tế bào lai sẽ mang các phân đoạn nhiễm sắc thể người khác nhau. Các dòng tế bào lai chiếu xạ sau đó sẽ được "sàng lọc" để tìm ra sự có mặt hay vắng mặt các đoạn trình tự gen người. Các locut có tính liên kết chặt hiếm khi bị tách ra trong các dòng tế bào lai khác nhau và có xu hướng cùng liên kết hoặc cùng bị mất đi trong các dòng tế bào lai. Ngược lại, các locut nằm xa nhau có xu hướng liên kết với các nhiễm sắc thể tế bào chủ hoặc bị mất đi một cách độc lập. Nhờ vậy, người ta có thể xác định vị trí các gen trên bản đồ vật lý. Trong phương pháp này, liều lượng chiếu xạ càng cao, càng có nhiều nhiễm sắc thể bị đứt gãy và bản đồ vật lý càng chi tiết.

Phương pháp nghiên cứu này cũng có thể được dùng để xây dựng bản đồ của toàn bộ hệ gen hay của từng nhiễm sắc thể riêng rẽ. Việc xây dựng bản đồ từng nhiễm sắc thể có thể thực hiện bằng việc sử dụng các tế bào lai soma [người × chuột] chỉ chứa một nhiễm sắc thể người duy nhất làm nguồn vật liệu để tiến hành tạo con lai chiếu xạ. ADN có nguồn gốc từ các dòng con lai chiếu xạ riêng rẽ được sàng lọc nhờ kỹ thuật thẩm tách Southern (sử dụng các mẫu dò) hoặc bằng kỹ thuật PCR (sử dụng các cặp mồi đặc hiệu cho dấu chuẩn ADN). Trong phương pháp này, khoảng cách vật lý giữa các dấu chuẩn có thể tính từ tần số ADN bị đứt gãy do chiếu xạ, vì tần số này tỷ lệ thuận với khoảng cách vật lý giữa các dấu chuẩn.

Một dạng bản đồ vật lý khác được thiết lập dựa trên việc sắp xếp các phân đoạn nhiễm sắc thể người sau khi tách dòng thành các trình tự nằm gối lên nhau liên tiếp, gọi là **contigs** (*contig*uous overlapping fragments, *Hình 6.12*). Kỹ thuật này có thể thực hiện đối với các phân đoạn nhiễm sắc thể lớn để lập bản đồ kích thước lớn hoặc đối với các vùng nhỏ hơn trong các nghiên cứu chi tiết. Để lập bản đồ gen các phân đoạn lớn, vector tách dòng loại nhiễm sắc thể nhân tạo nấm men YAC (*y*east *a*rtificial *c*hromosome) thường được sử dụng để mang các đoạn trình tự ADN có kích thước lên đến 2Mbp. Tuy vậy, vector này có một số nhược điểm, như kém bền, hoặc có thể làm mất một hoặc một số đoạn của đoạn nhiễm sắc thể tái tổ hợp; có thể tạo ra dạng khảm do các đoạn nhiễm sắc thể từ các phần khác nhau của hệ gen có thể gắn vào cùng một vector. Các contigs từ các đoạn trình tự ADN lớn ở tất cả các nhiễm sắc thể người đã được thiết lập và chuyển vào vector YAC. Các trình tự contigs sau đó được định vị trên bản đồ khung bằng việc xác định các vị trí STS (*s*equence *t*agged *s*ites). Một ví dụ cho phương pháp nghiên cứu này là việc xác định vị trí các chỉ thị contigs nằm dọc nhiễm sắc thể số 21.

Khác với YAC, vector cosmid có thể mang những đoạn ADN kích thước nhỏ hơn (30 - 45kb) nhưng bền vững hơn. Tuy vậy, do nó chỉ có thể mang những đoạn ADN nhỏ, nên việc xây dựng các trình tự contigs sẽ cần nhiều thời gian và công sức hơn. Thông thường, người ta phải phối hợp hai trình tự contigs để có thể tìm một vùng đặc hiệu trên một gen mới. Vì vậy, nếu như người ta sử dụng YAC để xây dựng bản đồ sơ bộ, thì phần chi tiết của bản đồ thường

được bổ sung bằng các chỉ thị contigs thu được từ các vector tách dòng cosmid (xem mô tả trên *Hình 6.12*).



Hình 6.12. **Các đoạn trình tự nằm gối lên nhau (contigs) được tách dòng bởi vector YAC cung cấp thông tin sơ bộ cho các phân tích tiếp theo. Nếu một gen có cả ở YAC1 và YAC3 thì nó phải nằm ở vùng gối lên nhau của hai dòng vector tách dòng này. Vùng trình tự nằm gối lên nhau này tiếp tục được phân tích bằng cách tách dòng theo nguyên tắc các trình tự nằm gối lên nhau vào vector cosmid. Nhờ vậy, các đoạn contigs mới có độ phân giải cao hơn sẽ được tạo ra và cho phép phân tích chi tiết và chính xác vị trí của gen cần quan tâm**

Gần đây, có hai loại vector tách dòng mới được phát triển, đó là nhiễm sắc thể nhân tạo vi khuẩn BAC (*b*acterial *a*rtificial *c*hromosome) và nhiễm sắc thể nhân tạo P1 (PAC = P1 *a*rtificial *c*hromosome). Những vector này có độ bền cao như cosmid, nhưng khả năng tái tổ hợp dạng khảm thấp, hơn nữa có thể mang các đoạn ADN ngoại lai có kích thước đến 300bp.

Phương pháp lai huỳnh quang *in situ*

Phương pháp lai huỳnh quang *in situ* có vai trò quan trọng trong các nghiên cứu lập bản đồ hệ gen người. Trong phương pháp này, người ta tiến hành lai các mẫu dò đánh dấu phóng xạ với các đoạn ADN đích trên các mẫu vật phẩm tế bào, chứ không phải trên các màng lai. Các mẫu dò có thể là các trình tự ADN hệ gen hoặc cADN sau khi được lai với các nhiễm sắc thể sẽ giúp xác định vị trí của các trình tự tương ứng của chúng trên các nhiễm sắc thể. Trước đây, người ta thường sử dụng các mẫu dò đánh dấu phóng xạ, nhưng phương pháp này có độ phân giải thấp. Hiện nay, các mẫu dò thường được đánh dấu huỳnh quang hoặc bằng việc sử dụng các base nitơ cải tiến có thể phát hiện bằng các kháng thể phát huỳnh quang.

Sắp xếp vị trí các gen trên bản đồ

Các bản đồ khung chứa hầu như toàn bộ các dấu chuẩn đa hình từ các vùng ADN không mã hóa. Trong thực tế, khi hệ gen được giải mã hoàn chỉnh, thì người ta có thể xác định một cách trực tiếp vị trí của tất cả các gen, nhưng hiện nay một trong những nhiệm vụ chủ yếu đối với dự án giải mã hệ gen người là chỉ ra vị trí của các gen. Điều này được thực hiện bởi một số phương pháp. Phương pháp thứ nhất gọi là **tách dòng vị trí** (positional cloning). Phương pháp này phụ thuộc vào việc xác định vị trí của gen liên quan đến một hội chứng di truyền nào đó có ở gần hay không một locut đã được xác định trên bản đồ khung. Nếu gen đó nằm đủ gần locut dấu chuẩn (dưới 2Mb), thì vùng gen đó sẽ được phân tích tìm sự có mặt của các trình tự mã hóa, và các trình tự tìm thấy sẽ được đánh giá xem liệu chúng có liên quan đến hội chứng di truyền đó hay không. Các gen đã được phân lập từ mARN và cADN có thể được lập bản đồ nhờ việc sàng lọc các vector YAC trước đó đã được

xác định vị trí với các contigs. Một phương pháp khác là người ta sử dụng các con lai chiếu xạ để xác định locut của gen tương ứng.

Đến nay, mới chỉ có một tỷ lệ nhỏ các gen trong hệ gen người được biết đến. Tuy vậy, hiện nay người ta có thể lập bản đồ của kể cả các gen chưa được xác định. Quá trình này phụ thuộc vào việc phân lập được các trình tự biểu hiện **EST (*expressed sequence tags*)** còn được gọi là **bản đồ phiên mã (transcript mapping)**. EST là các trình tự đặc thù duy nhất của các phân tử mARN chưa xác định được dùng để thiết kế các đoạn mồi nhằm sử dụng phản ứng PCR để tìm các trình tự mã hóa tương ứng có mặt trong ADN hệ gen. Quy trình xác định này bắt đầu bằng việc thiết lập thư viện cADN từ các phân tử mARN chứa đuôi poly(A). mARN được thu nhận từ nhiều loại mô khác nhau để làm tăng số loại phân tử mARN được dùng để tạo ra các cADN. Các phân tử cADN sau đó được kiểm tra để loại bỏ các bản sao mARN giống nhau. Các phân tử cADN tương ứng thu được sau đó sẽ được giải mã trình tự để xác định các trình tự đoạn mồi đặc trưng duy nhất cho đoạn trình tự gen tương ứng. Các đoạn mồi này tiếp tục được dùng để sàng lọc các con lai chiếu xạ, cũng như các contigs trên nhiễm sắc thể nhân tạo nấm men. Phương pháp nghiên cứu trên đây rất hiệu quả do hơn một nửa số gen của người hiện nay được mã hóa ở dạng cADN. Mục đích của phương pháp là xác định các chỉ thị EST cho tất cả các vùng của hệ gen, nhờ vậy bất cứ khi nào các gen liên quan đến một hội chứng di truyền được định vị trên một nhiễm sắc thể nào đó, thì hàng loạt EST sẽ được xác định cho vùng ADN lân cận.

VI.F DI TRUYỀN HỌC VÀ KHOA HỌC HÌNH SỰ

Sự tương quan đặc trưng cá thể

Tính đặc trưng của các cá thể (trong các loài sinh sản hữu tính) có thể được sử dụng để xác định từng cá thể riêng biệt, hoặc mọi cấu trúc tế bào là thành phần của cá thể đó. Bằng việc sử dụng phản ứng chuỗi trùng hợp PCR, từ một vài tế bào của một cơ thể nào đó, chúng ta có thể xây dựng được một "hồ sơ" di truyền học của cơ thể đó. Nhờ vậy, chúng ta có thể có thể biết được sự đặc trưng di truyền của từng cá thể một cách chắc chắn. Mọi mô trong cơ thể đều chứa cấu trúc ADN nguyên vẹn của cá thể đó và có thể sử dụng làm nguồn mẫu cung cấp ADN cho quá trình phân tích, chẳng hạn như tinh dịch (trong các vụ án hiếp dâm, cưỡng hiếp, ...), máu, lông, tóc, các mảnh da, các tế bào biểu mô miệng (trong tuyến nước bọt), hay xương của các bộ hài cốt.

Khoa học pháp y có thể so sánh các đặc trưng di truyền của các mẫu sinh học được tìm thấy ở hiện trường các vụ án hình sự để xác định thủ phạm hay loại bỏ nghi can vô tội. Mọi mẫu sinh học thu được từ một cá thể sẽ trùng hợp với "dấu vân tay di truyền" của cá thể đó về mọi đặc tính di truyền. Ý nghĩa quan trọng nhất của kỹ thuật này là xác định được và loại bỏ các nghi can vô tội. Chỉ cần một đặc tính di truyền không trùng với "dấu vân tay di truyền" của tội phạm là đủ để xác định được nghi can vô tội. Ngược lại, nếu các bằng chứng về dấu chuẩn di truyền là trùng khớp thì có thể khẳng định dấu vết để lại hiện trường chính là của nghi can. Lúc đó, tòa án có thể yêu cầu nghi can giải thích về sự liên quan của mình đến hiện trường vụ án.

Sự tương quan di truyền giữa "bố, mẹ" và "con cái" cũng có thể được dùng để khẳng định quan hệ huyết thống của họ. Đối với mỗi gen, đứa trẻ phải nhận được một alen từ bố và alen còn lại từ mẹ, do đó một nửa kiểu gen của đứa trẻ sẽ giống bố và nửa còn lại sẽ giống mẹ. Trên nguyên tắc đó, phép phân tích di truyền có thể được sử dụng làm cơ sở để khẳng định hoặc loại trừ một nghi vấn về quan hệ huyết thống giữa một người cha và đứa con. Chẳng hạn, trong những trường hợp "nhầm lẫn" trẻ sơ sinh trong nhà hộ sinh, người ta có thể sử dụng kỹ thuật này để xác định được chính xác bố, mẹ của đứa trẻ. Ngoài ra, kỹ thuật phân tích di truyền cũng có thể sử dụng trong khoa học hình sự trên các đối tượng sinh vật khác, ví dụ như để xác định những con ngựa, trâu, bò ... bị mất trộm.

Trong thực tế một số vụ án hình sự, tội phạm đã bị phát hiện nhờ việc phân tích dấu vân tay di truyền trên các đối tượng sinh học không phải con người. Chẳng hạn, trong một vụ án hạt của một loài cây nằm sát ngay cạnh nạn nhân ở hiện trường vụ án đã được tìm thấy trên xe của nghi can, hoặc trong trường hợp khác lông mèo để tại hiện trường trùng khớp với lông của con mèo của nghi can. Trong hai trường hợp như trên, sai số kết luận nghi can có liên quan đến hiện trường vụ án thấp hơn $1/10^6$. Vì vậy, trong những trường hợp này, nếu bằng chứng ADN đã được khẳng định thì các nghi can hầu như chắc chắn có liên quan đến hiện trường vụ án.

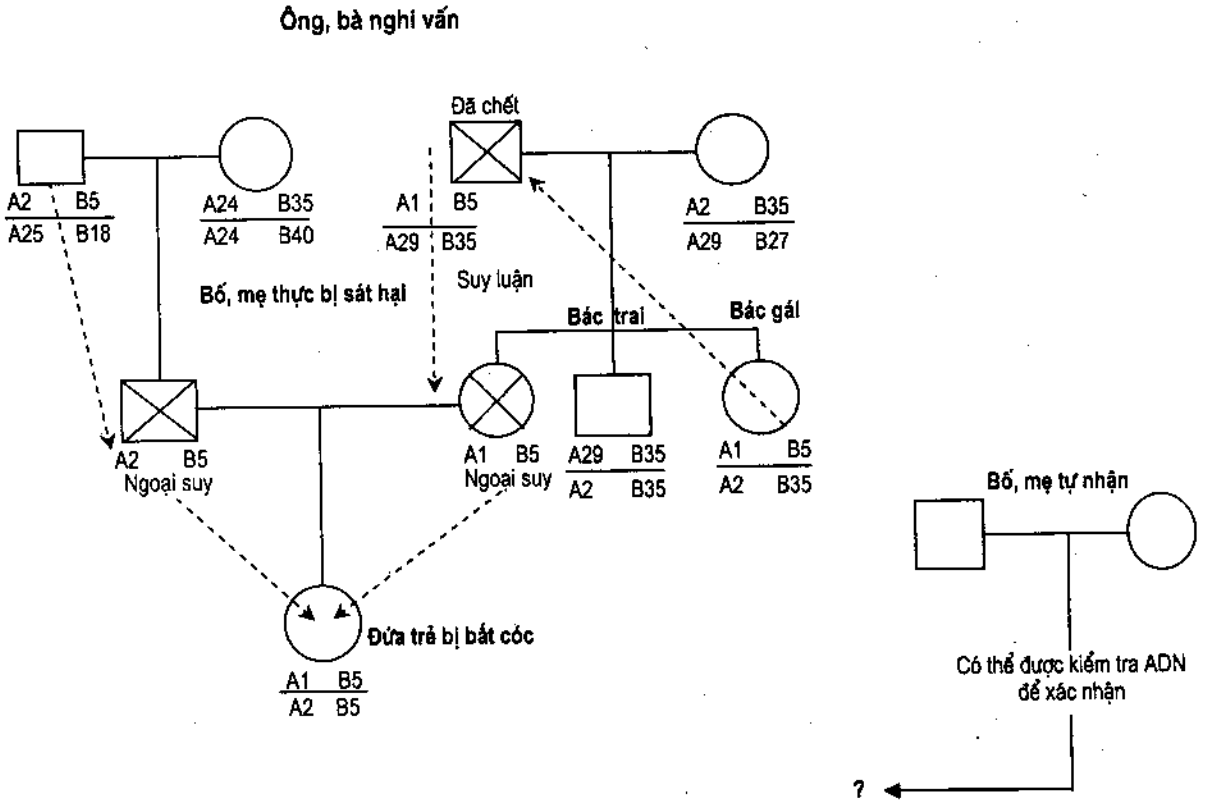
So sánh mẫu protein

Trình tự của các amino acid trong phân tử protein được xác định bởi cơ chế di truyền, và có thể kiểm tra trực tiếp bằng phương pháp phân tích protein, hoặc gián tiếp thông qua sự thay đổi về tính trạng mà protein đó quy định. Chẳng hạn, tính trạng màu mắt xanh là một tính trạng do gen quy định. Việc phân tích tính trạng này ít có ý nghĩa ở vùng Xcăngđinavia vì phần lớn người dân ở đây có màu mắt xanh, nhưng nó sẽ rất hữu dụng khi được dùng để phân tích ở một đất nước mà màu mắt nâu, hoặc mắt đen là phổ biến. Như vậy, việc phân tích một tính trạng trong điều tra hình sự có hiệu quả hay không còn phụ thuộc vào mức độ phổ biến của tính trạng đó. Các thí nghiệm phân tích đầu tiên thường được tập trung vào các phân tử protein đặc thù bằng việc sử dụng nguyên lý miễn dịch. **Các kháng thể** được hệ thống miễn dịch tạo ra để phản ứng lại các phân tử ngoại lai. Chúng ta biết rằng các loài động vật có vú, như thỏ, cừu, dê, ngựa khi được truyền máu của người, chúng sẽ tạo ra kháng huyết thanh chứa các kháng thể có khả năng bám vào các tế bào hồng cầu của máu người. Các kháng huyết thanh này có thể giúp phân biệt các nhóm máu khác nhau như A, B, và O; cũng như một số loại nhóm máu khác như M-MN-N, tùy theo kiểu nhóm máu đầu tiên được sử dụng để truyền vào con vật. Các kháng thể này cũng có thể được dùng để xác định loại nhóm máu chỉ cần dựa trên một giọt máu duy nhất. Nếu như vết máu giống của nạn nhân được tìm thấy trên cơ thể, quần áo và đồ dùng của nghi can, hoặc ngược lại, thì có thể xác định có hay không sự tiếp xúc của nghi can với nạn nhân. Các alen càng hiếm gặp, thì càng hữu dụng cho quá trình loại trừ nghi can. Chẳng hạn nhóm máu N chỉ chiếm 1,1% trong số những người Ấn Độ dòng Navaho, nhưng chiếm tới 21% trong quần thể người dân vùng Cápca, 67% trong những người thổ dân Châu Úc. Tính trạng này có thể sử dụng để loại trừ 98,9% người Navaho, 80% người dân Cápca, nhưng chỉ loại trừ được 33% người thổ dân Úc.

Các alen đặc trưng mô thường có mức độ biến dị lớn, và vì vậy thường được sử dụng phổ biến để xác định huyết thống của những đứa trẻ mất tích. Chẳng hạn, ở Achentina, do bất

đông chính trị, có những đứa trẻ bị bắt cóc và bố, mẹ chúng bị sát hại. Khi người ta tìm được đứa trẻ, các alen đặc trưng mô của chúng sẽ được so sánh với các cặp ông, bà nghi vấn của chúng. Việc này giúp xác định quan hệ huyết thống của đứa trẻ là phù hợp với cặp ông, bà nào (xem minh họa *Hình 6.13*).

Một kỹ thuật khác được gọi là **điện di protein**, quá trình này giống với điện di các acid nucleic. Một đột biến có thể làm thay đổi một amino acid duy nhất trong một phân tử protein, nhưng điều đó cũng đủ làm biến đổi điện tích và khối lượng của phân tử protein đó, và vì vậy làm thay đổi tốc độ di chuyển của phân tử protein trên trường điện di. Những đột biến kiểu như vậy rất phổ biến. Một lần nữa, nguyên tắc này được sử dụng để so sánh mẫu tìm thấy ở hiện trường vụ án và đối tượng bị tình nghi.



Hình 6.13. **Phả hệ của các alen HLA (alen đặc trưng mô) của ông, bà và đứa trẻ bị bắt cóc. Mũi tên - - → truy nguyên nguồn gốc các alen của đứa trẻ với các thành viên khác trong gia đình, khẳng định sự có mặt các alen của đứa trẻ trong gia đình mà bố, mẹ đã chết, cho thấy 99,9% mối quan hệ huyết thống đã xác định là đúng.**

So sánh ADN

Mọi cơ thể được hình thành nhờ quá trình sinh sản hữu tính đều có một kiểu trình tự ADN đặc trưng duy nhất. Nói cách khác, mỗi cá thể có một kiểu gen đặc thù. Khi phân tử ADN được thu thập từ mô còn để lại tại hiện trường vụ án trùng khớp với trình tự ADN của một nghi can, thì có thể khẳng định rằng mẫu mô để lại tại hiện trường là của nghi can đó (tất nhiên với điều kiện trình tự so sánh là đủ lớn). Việc phát triển các kỹ thuật thao tác ADN, đặc biệt việc sử dụng các enzym giới hạn ngày nay cho phép tiến hành các phép so sánh các trình tự ADN một cách dễ dàng.

Các kỹ thuật so sánh ADN được sử dụng trong khoa học hình sự về cơ bản đều dựa trên việc phân tích điện di các phân đoạn ADN. Các phân đoạn ADN đầu tiên được cắt ra từ ADN tổng số của hệ gen nhờ các enzym giới hạn. **Các enzym giới hạn** xúc tác việc cắt phân tử ADN tại những trình tự nhất định (được gọi là các vị trí giới hạn) vốn xuất hiện tương đối ngẫu nhiên trong khắp phân tử ADN hệ gen (ví dụ: enzym *EcoR*I cắt ở trình tự GAATTC thường xuất hiện với tần số trung bình là từ 2000 đến 5000 cặp base nitơ tương ứng đối với các loài có tỷ số AT: GC cao và thấp). Các đoạn ADN được cắt bởi enzym giới hạn được gọi là các đoạn cắt giới hạn.

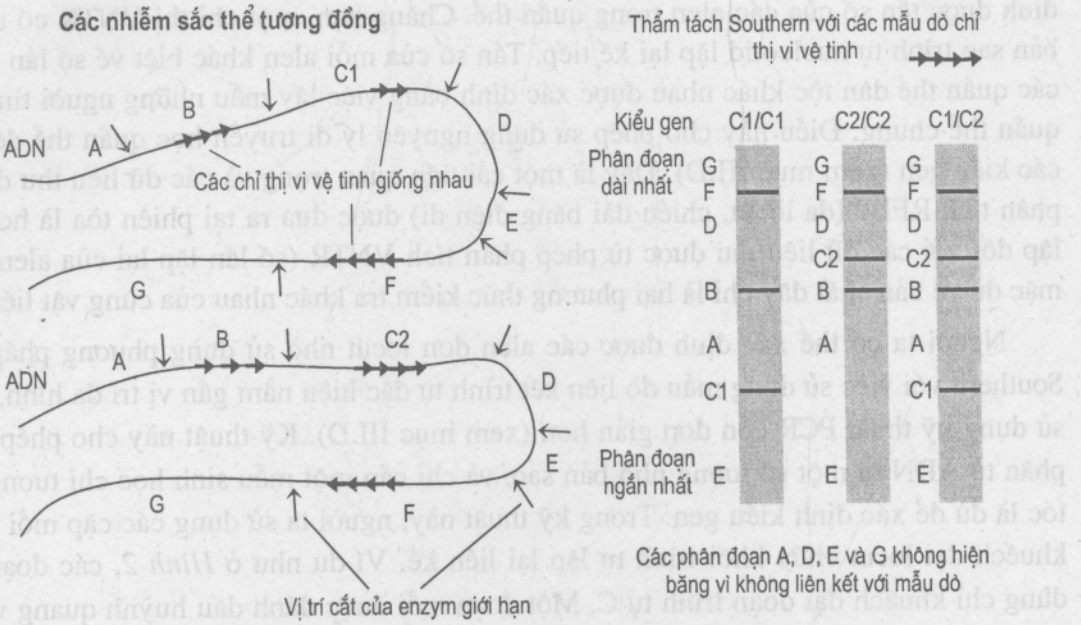
Nếu các trình tự ADN luôn giống hệt nhau, thì độ dài các đoạn cắt giới hạn tương ứng của tất cả mọi người sẽ giống nhau, nhưng rất may chúng thường khác nhau. Các đột biến có thể làm thay đổi độ dài của một đoạn cắt giới hạn bằng 1 trong 2 cách sau: (i) làm thay đổi vị trí cắt giới hạn. Đột biến có thể tạo ra một vị trí cắt giới hạn mới bên trong đoạn cắt giới hạn cũ, do đó đoạn cắt giới hạn mới trở nên ngắn hơn (ví dụ: trình tự GAAATC có thể đột biến thành GAATTC) hoặc đột biến cũng có thể làm biến đổi một vị trí cắt giới hạn cũ thành một trình tự khác (không phải trình tự cắt giới hạn), lúc đó hai đoạn cắt giới hạn ở hai đầu vị trí cắt sẽ nối với nhau thành một đoạn cắt giới hạn dài hơn; (ii) các cặp base nitơ có thể được gắn thêm vào hoặc mất đi, làm cho các đoạn cắt giới hạn có thể dài ra hoặc ngắn lại. Hiện tượng này xảy ra phổ biến do sự thay đổi số bản sao có mặt trong khối **các trình tự lặp lại liền kề** vốn không bền và thường có tính đa hình cao trong các quần thể. Những biến dị này được gọi là chỉ thị **biến đổi số trình tự lặp lại liền kề VNTR** (*v*ariable *n*umber *t*andem *r*epeats). Nếu các đơn vị lặp lại ngắn hơn 6 cặp base nitơ, thì chúng được gọi là chỉ thị vi vệ tinh (microsatellite); nếu gồm từ 6 - 30 cặp base nitơ thì được gọi là chỉ thị tiểu vệ tinh (minisatellite); còn nếu đơn vị lặp lại dài hơn, thì được gọi chung là chỉ thị vệ tinh VNTR.

Phần lớn các biến dị được sử dụng trong khoa học hình sự nằm trong các locut VNTR. Vì mức độ đa hình của chúng rất lớn, nên khả năng có hai người có kiểu hình chỉ thị VNTR giống hệt nhau là rất thấp. Có hai kỹ thuật cơ bản được các cán bộ khoa học hình sự thường sử dụng để phân tích các chỉ thị VNTR. Trong kỹ thuật kinh điển, tất cả các đoạn ADN chứa chỉ thị vệ tinh đều được phân tích. Trong trường hợp này, vì dọc hệ gen có nhiều đoạn chứa trình tự vệ tinh, nên hệ thống này gọi là hệ thống **đa locut** (multilocus). Hệ thống này tạo ra một kiểu dấu vân tay ADN đặc thù bởi phổ kết hợp của nhiều băng điện di khác nhau. Tuy vậy, gần đây kỹ thuật được ưa chuộng hơn là sử dụng phản ứng PCR với một số mồi đặc hiệu để phát hiện ra các chỉ thị tiểu vệ tinh ở một số vị trí **đơn locut** (single locus) nhất định. Đó là bởi vì mỗi một cá thể chỉ có hai alen tại một locut và nhờ vậy việc phân tích kết quả điện di sẽ dễ dàng hơn.

RFLP: dấu vân tay di truyền

Chỉ thị ADN đầu tiên được sử dụng trong khoa học hình sự là chỉ thị **đa hình độ dài đoạn phân cắt giới hạn RFLP** (*r*estriction *f*ragment *l*ength *p*olymorphism). Các đoạn cắt giới hạn được phân tách trên gel điện di dựa vào kích thước (chiều dài) của chúng. Việc cắt phân tử ADN tổng số của một cơ thể đa bào vốn có hệ gen lớn sẽ chứa đựng một tập hợp các đoạn cắt giới hạn có một chiều dài khác nhau không lớn, vì vậy việc phân tách và phân biệt từng đoạn cắt giới hạn gặp khó khăn. Vì vậy, phải có phương pháp để chọn ra và

phân tích các băng đặc thù. Quá trình này được gọi là kỹ thuật **thẩm tách Southern** sử dụng **mẫu dò đa locut**. Đây là những trình tự xuất hiện ở nhiều vị trí trong hệ gen, một cách lý tưởng, chúng có khả năng tạo ra được nhiều đoạn băng điện di phân tách trên trường điện di (*Hình 6.14*). Nếu việc cắt hai mẫu ADN được chuẩn bị tốt từ cùng một cá thể được chạy điện di trên hai giếng gần nhau của bản gel (nhưng không làm nhiễm chéo) thì người ta sẽ thu được hai dải điện di giống hệt nhau về số lượng và độ đậm, nhạt của các băng điện di. Tuy vậy, nếu hai mẫu này được chạy trên hai bản gel điện di khác nhau, thì việc so sánh gặp khó khăn hơn, và nếu vị trí của các băng trên gel được đo một cách tương đối so với các băng chuẩn đã biết trước kích thước thì cần phải cân nhắc sai số do quá trình đo gây ra, và có thể không ghi nhận được độ đậm nhạt chính xác của từng băng. Những dữ liệu đó sẽ ít thuyết phục hơn trong các phiên tòa và vì vậy người ta có thể đặt câu hỏi về độ tin cậy.



Hình 6.14. **Kỹ thuật phân tích dấu vân tay di truyền sử dụng mẫu dò đa locut để phân tích thẩm tách Sourthern các phân đoạn cắt bằng enzym giới hạn cho thấy sự biến động về chiều dài của phân đoạn C. Mũi tên được dùng để minh hoạ vị trí và chiều dài đoạn chỉ thị vệ tinh trong phân tử ADN. Mỗi một locut có một hoặc hai băng trên gel điện di (tùy thuộc vào trạng thái đồng hợp tử hay dị hợp tử).**

Nhược điểm của chỉ thị RFLP:

Các chỉ thị RFLP hiện nay không còn được dùng phổ biến như trước trong các nghiên cứu pháp y vì những lý do sau: (i) kỹ thuật thường đòi hỏi một lượng lớn mẫu ADN đủ để có thể tiến hành lai với các mẫu dò trong phương pháp thẩm tách Southern; (ii) gặp khó khăn trong việc phân tích các dữ liệu thu được liên quan đến việc tính xác suất có thể trùng hợp ngẫu nhiên giữa hai mẫu của các cá thể khác nhau. Hạn chế về độ chính xác trong phép đo chiều dài các đoạn giới hạn trong các băng riêng lẻ, bao gồm cả sai số do các bản gel khác nhau gây ra. Trong trường hợp đó, không thể khẳng định các băng đại diện cho các alen khác nhau của cùng một locut, vì vậy các nguyên lý kinh điển của di truyền học quần thể không áp dụng

được để tính tần số một kiểu gen nhất định. Khó khăn này được giải quyết nhờ việc sử dụng các mẫu dò đơn locut và các kỹ thuật dựa trên PCR.

VNTR: hồ sơ di truyền

Các điều tra hình sự thường sử dụng các biến dị về chiều dài tại một số vị trí đơn locut chứa một chỉ thị VNTR. Khoảng 0,5% hệ gen người chứa các trình tự lặp lại dCA (= dTG), trình tự lặp lại dCT = dAG chiếm khoảng 0,2%. Các trình tự này là các trình tự nucleotid lặp lại gồm 2 nucleotid, nhưng các trình tự lặp lại gồm 4 nucleotid (như CAGA) thường được ưa dùng hơn trong nghiên cứu pháp y vì việc mất đi một trình tự gây ra một biến đổi rõ rệt hơn về chiều dài, hơn nữa các trình tự này ổn định hơn trong quá trình khuếch đại bằng PCR. Các chỉ thị VNTR thường biểu hiện tính đa hình lớn hơn khi phân tích bằng các chỉ thị đa locut, nhưng đối với khoa học hình sự, chúng thường được kiểm tra ở những locut đơn lẻ bằng **mẫu dò đơn locut**, và qua đó xác định được số lượng chính xác bản sao của các trình tự lặp lại liền kề có trong các alen đặc thù. Do trình tự các nucleotid được xác định, nên người ta có thể xác định được tần số của các alen trong quần thể. Chẳng hạn, một chỉ thị VNTR có từ 12 đến 23 bản sao trình tự nucleotid lặp lại kế tiếp. Tần số của mỗi alen khác biệt về số lần lặp lại trong các quần thể dân tộc khác nhau được xác định bằng việc lấy mẫu những người tình nguyện từ quần thể chung. Điều này cho phép sử dụng nguyên lý di truyền học quần thể để tính tần số các kiểu gen (xem mục III.D). Đây là một cải tiến quan trọng và các dữ liệu thu được từ phép phân tích RFLP (đa locut, chiều dài băng điện di) được đưa ra tại phiên tòa là hoàn toàn độc lập đối với các dữ liệu thu được từ phép phân tích VNTR (số lần lặp lại của alen đơn locut), mặc dù về bản chất đây chỉ là hai phương thức kiểm tra khác nhau của cùng vật liệu di truyền.

Người ta có thể xác định được các alen đơn locut nhờ sử dụng phương pháp thẩm tách Southern với việc sử dụng mẫu dò liên kết trình tự đặc hiệu nằm gần vị trí đa hình, nhưng việc sử dụng kỹ thuật PCR còn đơn giản hơn (xem mục III.D). Kỹ thuật này cho phép khuếch đại phân tử ADN từ một số lượng nhỏ bản sao, và chỉ cần một mẫu sinh học chỉ tương đương sợi tóc là đủ để xác định kiểu gen. Trong kỹ thuật này, người ta sử dụng các cặp mồi đặc hiệu để khuếch đại locut chứa khối trình tự lặp lại liền kề. Ví dụ như ở *Hình 2*, các đoạn mồi được dùng chỉ khuếch đại đoạn trình tự C. Một đoạn mồi được đánh dấu huỳnh quang và chiều dài băng được xác định nhờ một thiết bị phân tích gel tự động hoạt động theo nguyên tắc giống với giải mã trình tự ADN tự động. Một cá thể dị hợp tử sẽ tạo ra hai băng cho một locut. Trong thực tế, người ta thường sử dụng đồng thời 4 mồi cùng lúc để có thể phát hiện được 4 locut cùng một lúc, nhờ vậy thu được 8 băng điện di từ tất cả 4 locut của một cơ thể dị hợp tử lưỡng bội. Phương pháp này gọi là **multiplexed PCR**. Các locut thường được chọn từ các nhiễm sắc thể khác nhau để tránh hiện tượng liên kết và đảm bảo việc phân loại ngẫu nhiên. Kiểu phân tích này thường cho thấy rõ mức độ đa hình cao trong quần thể và do sự khác biệt rõ rệt về chiều dài trình tự giữa các locut nên việc phân biệt được dễ dàng, ngoài ra việc đánh dấu huỳnh quang với các màu khác nhau đối với mỗi loại mồi cho phép phân biệt được từng loại locut.

Phương pháp phân tích dữ liệu

Chỉ một sự khác biệt giữa hai mẫu so sánh là căn cứ xác minh hai mẫu đó có nguồn gốc khác nhau. Nếu kết quả phân tích của hai mẫu trùng khít lên nhau, thì

xác suất để có thể thu được hai mẫu giống hệt nhau (trên cơ sở phép phân tích) ngẫu nhiên có thể có trong quần thể sẽ được tính toán. Trong trường hợp này, việc phân tích "vân tay" ADN tương đối khó do việc xác định vị trí chính xác của các băng điện di đôi khi gặp khó khăn (Tuy vậy, trong thực tế, hai mẫu có nguồn gốc khác nhau thường luôn có biểu hiện khác nhau). Việc phân tích tính đa hình của VNTR cho phép xác định mỗi alen dựa vào số lần lặp lại của nó trong các dấu chuẩn vệ tinh. Qua đó, có thể xác định được tần số của mỗi alen trong một quần thể, và tính được xác suất để hai mẫu phân tích có thể giống hệt nhau một cách ngẫu nhiên. Một mức độ nội phối nào đó được ước lượng, và việc tính toán sẽ được hiệu dịch để làm giảm tối đa nguy cơ có những kết luận nhầm lẫn khi phân tích dấu chuẩn di truyền.

Ở một số quốc gia, các thẩm phán được đào tạo nhằm cố gắng xây dựng hiện trường từ các dẫn liệu thu được, nhưng ở Anh và Mỹ, người ta sử dụng một hệ thống khác dựa trên việc thuyết phục đoàn hội thẩm về việc đúng đắn của một bên trong vụ án, và sự sai trái của bên còn lại. Đoàn hội thẩm thường gồm 12 hội thẩm viên (là những người không có chuyên môn phân tích các chứng cứ) và quan tòa lúc đó chỉ có vai trò dựa vào các chứng cứ để tuyên án, chứ không phải điều tra vụ án. Điều này dễ dẫn đến những phán quyết sai khi sử dụng các chứng cứ nhạy cảm, và vì vậy một số vụ án đầu tiên sử dụng các chứng cứ dấu vân tay di truyền đã bị loại bỏ do quan tòa không chắc chắn về cơ sở khoa học của các bằng chứng khi ủy viên công tố cương quyết khẳng định những chứng cứ đó không đảm bảo tính chính xác. Vì những nguyên nhân đó, các chứng cứ pháp y phải được trình bày rất cẩn thận theo một quy trình khoa học, chặt chẽ để tránh bị chỉ trích. Cho đến nay, đã có những vụ án mà các bằng chứng khoa học bị lãng quên. Vua hề Charlie Chaplin từng bị một nữ diễn viên tố cáo là bố đứa con của cô ta. Cuối cùng Chaplin bị kết án mặc dù các bằng chứng về nhóm máu khẳng định anh ta không thể là bố của đứa trẻ. Ngược lại, có những vụ án mà nghi can bị kết tội do kỹ thuật viên làm tràn mẫu ADN của nghi can vào giếng bên cạnh chứa mẫu thu thập tại hiện trường vụ án, tạo ra các băng điện di giống hệt nhau sau nhiều lần phân tích cho kết quả âm tính. Ảnh hưởng của những lỗi kỹ thuật đó có thể được nhận biết dễ dàng bởi các kỹ thuật viên lành nghề và chuyên gia. Vì lý do nêu trên, việc trình bày rõ ràng các chứng cứ di truyền ở phiên tòa còn quan trọng hơn chính quá trình phân tích tại phòng thí nghiệm.

Các dấu vân tay di truyền được sử dụng trong khoa học hình sự theo nguyên tắc loại trừ. Theo đó, chỉ một sai khác dù nhỏ giữa mẫu thu tại hiện trường mà mẫu của nghi can là đủ để loại trừ nghi can. Nhưng vấn đề lại không phải việc có thể loại trừ nhầm nghi can, mà lại là xác suất có thể kết tội nhầm một nghi can do sự trùng hợp ngẫu nhiên. Ví dụ, nếu một người bị kết tội hãm hiếp do có kiểu hình điện di ADN trùng với tinh dịch để lại hiện trường. Nếu tần số của kiểu gen đó trong quần thể là 0,01 (1%) thì xác suất để người bị tình nghi có kiểu hình băng điện di trùng với tội phạm là 1/100. Nếu chỉ dựa trên chứng cứ đó để kết tội nghi can, thì xác suất kết tội nhầm sẽ là 1 người có trong 100 người vô tội. Tuy vậy, nếu tần số kiểu gen là 10^{-8} thì xác suất trùng kiểu gen ngẫu nhiên chỉ là 1 trong 10^{8}. Mà điều này thì có nghĩa là hầu như không thể có một người thứ hai có kiểu gen tương tự trên cùng một châu lục giống với kiểu gen của kẻ tội phạm.

Các nhà nghiên cứu dấu vân tay truyền thống thường ít quan tâm đến nguyên tắc xác suất. Họ thường chỉ đơn giản xác định xem các điểm trùng khớp giữa 2 dấu vân tay là đủ lớn hay

chưa để kết luận 2 dấu vân tay đó có cùng xuất phát từ một ngón tay hay không. Các nhà di truyền học chỉ thực sự dựa vào nguyên tắc xác suất kể từ khi có những thí nghiệm của Mendel. Trên cơ sở đó, người ta xác định xác suất để một mẫu có thể trùng khớp với mẫu thu ở hiện trường do nguyên tắc ngẫu nhiên. Chỉ dựa vào câu hỏi liên quan đến nguyên tắc xác suất, có những luật sư bào chữa đã cố gắng bảo vệ nghi can dựa vào những nghi ngờ về mặt xác suất. Các dữ liệu đa locut đòi hỏi việc đo vị trí các băng để tính chiều dài các đoạn trình tự. Nếu sử dụng một bản gel dài 200mm với độ phân giải chính xác là 1mm thì tối đa đo được vị trí của 200 băng. Các băng chỉ cách nhau 0,5mm sẽ nằm sát liền vào nhau, và vì vậy làm giảm thông tin hữu ích.

Bằng việc sử dụng các mẫu dò đơn locut, các alen riêng lẻ có thể được xác định. Người ta có thể xác định được tần số của chúng trong quần thể và sử dụng công thức của Định luật Hardy - Weinberg (xem mục IV.C) để có thể tính toán tương đối chính xác xác suất để hai mẫu không có liên quan trùng khớp nhau một cách ngẫu nhiên. Một vấn đề tồn tại là khả năng xuất hiện trong quần thể một hoặc một số quần thể phụ có quá trình nội phối. Điều này tỏ ra là căn cứ chống đối lớn nhất đối với các dẫn chứng sử dụng VNTR. Nếu một nghi can và mẫu thu ở hiện trường đều có tần số alen hiếm gặp là 0,001 thì chỉ khoảng 1 trong cứ 500 người có mang alen đó, nhưng nếu một trong bố hoặc mẹ của nghi can mang alen đó thì một nửa số chị, em và con cháu sẽ mang alen đó. Nếu gia đình đó sống lâu dài trong một cộng đồng nhỏ thì tần số bắt gặp alen đó trong các thế hệ con cháu sẽ lớn hơn 1/500 nhiều. Vì vậy, xác suất trùng hợp ngẫu nhiên sẽ tăng lên. Điều này có thể khắc phục bằng việc xác định các alen hiếm gặp ở tần số tối thiểu là 0,05. Lúc đó phiên tòa sẽ được trao xác suất không loại trừ được người vô tội nếu người đó có quan hệ gần gũi (ví dụ là anh, chị, em sau 1 – 2 thế hệ) với thủ phạm với xác suất từ 1/10.000 đến 1/1.000.000. Việc sử dụng các locut khác sẽ khẳng định lại điều này, nhưng nếu các dẫn chứng khác càng khẳng định nghi can, thì chứng cứ càng được bổ sung.

Các chủng tộc ít người cũng chính là các "quần thể phụ". Các alen ở locut WVA ở người chứa khoảng từ 12 đến 21 trình tự lặp lại. Các alen chứa 12 trình tự lặp lại chỉ bắt gặp ở một chủng tộc người duy nhất. Số alen gồm 15 trình tự lặp lại có tần số khoảng 0,09 ở Châu Á và ở vùng Cápca, và 0,2% ở vùng biển Caribê ở Châu Phi. Tần số cao nhất tương ứng với nhóm chủng tộc nghi can hay nói cách khác quần thể này có nhiều nguy cơ bị kết tội. Nếu không có các dẫn chứng khác liên quan đến nhóm chủng tộc, thì việc tính toán theo nguyên tắc di truyền quần thể có thể là chính xác hơn.

VI.G KỸ THUẬT DI TRUYỀN VÀ CÔNG NGHỆ SINH HỌC

Kỹ thuật di truyền và công nghệ sinh học

Hai thuật ngữ này được dùng để chỉ lĩnh vực nghiên cứu sử dụng các kỹ thuật ADN tái tổ hợp để chuyển gen (biến nạp gen) từ cơ thể (hoặc loài) này sang cơ thể (hoặc loài) khác. Kỹ thuật biến nạp gen có hai ứng dụng quan trọng: thứ nhất, nó có thể cho phép tạo ra một số lượng lớn protein mang hoạt tính sinh học. Trong số này bao gồm các protein được sử dụng làm thuốc chữa bệnh (*Bảng 6.2*) và

các protein hoặc enzym được sử dụng trong công nghiệp (*Bảng 6.3*). Thứ hai, quá trình chuyển gen được sử dụng để tạo ra các cơ thể mới (thực vật, động vật và vi sinh vật) mang các đặc tính mong muốn, chẳng hạn như cây trồng có khả năng kháng sâu hoặc thuốc trừ cỏ. Các ứng dụng này của kỹ thuật di truyền và công nghệ sinh học có tiềm năng thương mại lớn và đã góp phần hình thành và phát triển của ngành công nghiệp sinh học

Bảng 6.2. **Các protein tái tổ hợp có nguồn gốc người được dùng làm dược phẩm**

Tên protein	Ứng dụng điều trị bệnh
α1-antitrypsin	Phù thũng
Calcitonin	Còi xương
Chorionic gonadotropin	Vô sinh
Erythropoietin	Thiếu máu
Protein yếu tố VIII	Máu loãng (khó đông máu)
Insulin	Tiểu đường
Interferon (α, β, γ)	Bệnh do viêm nhiễm virus, ung thư
Interleukin	Ung thư
Protein hoạt hóa plasminogen mô	Huyết khối
Hormon sinh trưởng	Chậm lớn

Bảng 6.3. **Các enzym tái tổ hợp được sử dụng trong công nghiệp**

Tên protein	Lĩnh vực ứng dụng
Rennin	Sản xuất phomát
α-amylase	Sản xuất bia
Bromelain	Làm mềm thịt, làm trong nước giải khát
Catalase	Chất chống oxy hóa thực phẩm
Cellulase	Sản xuất glucose và cồn
Lipase	Sản xuất phomát
Protease	Làm thuốc tẩy

Sự biểu hiện của các protein tái tổ hợp

Một trong những mục đích cơ bản của kỹ thuật di truyền và công nghệ sinh học là sản xuất ra các loại protein sinh học hữu ích. Quá trình này là kết quả biểu hiện các gen được tách dòng trong các tế bào chủ khác nhau. Các tế bào chủ này có thể là các tế bào vi khuẩn, nấm men, động vật hay thực vật. Sự biểu hiện của các protein tái tổ hợp thường đòi hỏi một vector đặc biệt được gọi là **vector biểu hiện** chứa trình tự ADN điều khiển quá trình phiên mã và dịch mã của gen trong tế bào chủ. Trong vector biểu hiện có thể còn bao gồm một số trình tự đặc biệt khác, như trình tự đảm bảo tính ổn định của protein biểu hiện hoặc điều khiển tế bào tiết protein ra môi trường để dễ dàng

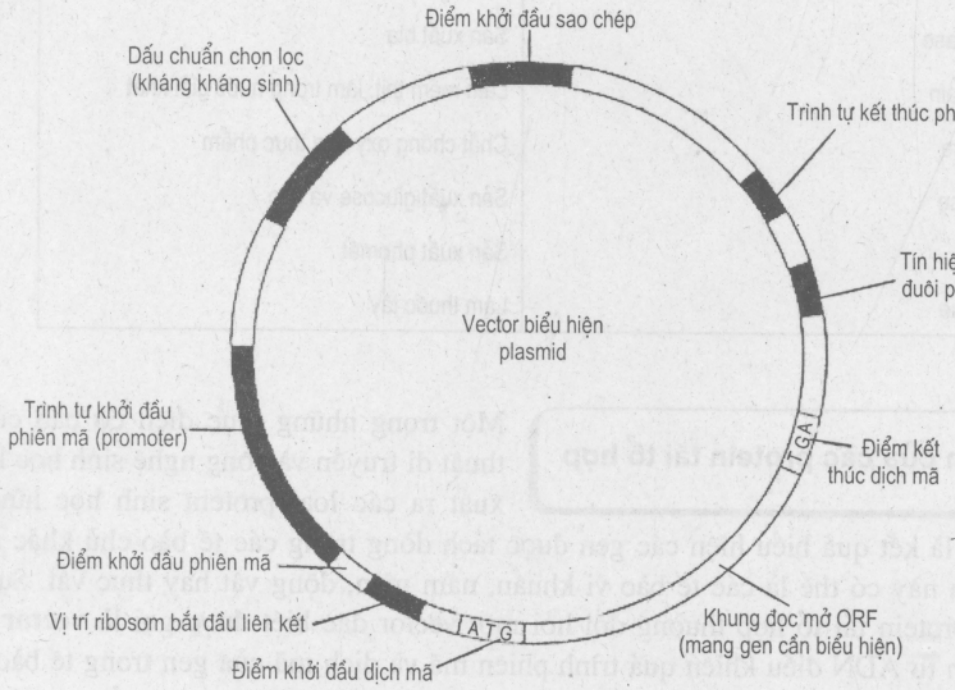
thu nhận. Nhằm đạt hiệu suất thu protein cao, người ta thường dùng vector plasmid do vector này có thể đồng thời xuất hiện nhiều bản sao trong tế bào chủ. Các điều kiện môi trường cần thiết để thu được hiệu suất biểu hiện của từng loại protein là khác nhau, vì vậy đối với từng loại protein thường phải nghiên cứu xác định các điều kiện tối ưu cho quá trình phiên mã và dịch mã. Cũng chính vì vậy, đến nay nhiều hệ thống biểu hiện protein khác nhau đã được tìm ra và phát triển.

Việc sử dụng công nghệ ADN tái tổ hợp để sản xuất protein đã tạo điều kiện cho sự hình thành các dạng cải biến của các phân tử protein tự nhiên. Quá trình này được gọi là **kỹ thuật cải biến protein**. Các trình tự gen được làm biến đổi nhờ phương pháp gọi là **gây đột biến *invitro*** để sau đó tạo ra các trình tự polypeptid với thành phần và vị trí các amino acid thay đổi so với dạng nguyên thủy, dẫn đến sự thay đổi các đặc tính và thuộc tính của protein. Kỹ thuật cải biến protein cũng được dùng để tạo ra các dạng biến đổi của protein thông thường, chẳng hạn nhằm làm tăng tính bền nhiệt của protein, hoặc đối với enzym làm thay đổi tính đặc hiệu cơ chất của chúng.

Hệ thống biểu hiện protein ở vi khuẩn

Những hệ thống biểu hiện protein tái tổ hợp đầu tiên đều sử dụng tế bào vi khuẩn là tế bào chủ. Trong đó, vi khuẩn *E. coli* có khả năng sản xuất nhiều loại protein với tốc độ nhanh và chi phí thấp được sử dụng rộng rãi hơn cả. Cho đến nay, các hệ thống biểu hiện nhờ tế bào vi khuẩn vẫn có vai trò quan trọng và được sử dụng rộng rãi để sản xuất phần lớn các loại protein quan trọng ở quy mô công nghiệp (*Bảng 6.2* và *6.3*).

Các vector biểu hiện protein ở vi khuẩn có một số đặc điểm cơ bản như sau (*Hình 6.15*):



Hình 6.15. **Đặc điểm của các vector biểu hiện plasmid**

– Một trình tự khởi đầu phiên mã (promotor) nằm ở phía trước trình tự gen tách dòng có chức năng khởi đầu phiên mã. Việc chọn lọc trình tự khởi đầu phiên mã có ái lực cao với ARN polymerase sẽ đảm bảo cho hiệu suất phiên mã cao.

– Một trình tự là tín hiệu kết thúc phiên mã nằm phía sau trình tự gen tách dòng.

– Một trình tự gắn ribosom (còn gọi là trình tự Shine-Dalgarno) nằm phía sau vị trí(điểm) bắt đầu phiên mã. Mức độ tổng hợp protein phụ thuộc vào ái lực của ribosom đối với trình tự này.

– Trình tự gen được tách dòng thường có bộ ba khởi đầu phiên mã AUG ở đầu 3' và bộ ba kết thúc phiên mã ở đầu 5'. Nằm ở giữa là một khung đọc mở. Một điểm quan trọng là khung đọc mở đó phải đảm bảo để sản phẩm biểu hiện của gen là trình tự mong đợi của các amino acid. Các vector thường được sử dụng đồng thời để biểu hiện các trình tự gen tách dòng theo một số cách khác nhau, nhờ vậy một trong ba khung đọc có thể được biểu hiện tùy theo yêu cầu.

– Các vector cho phép sự kích hoạt phiên mã của gen tái tổ hợp. Trong khi đó, việc biểu hiện liên tục một protein nào đó có thể sẽ tiêu tốn một phần lớn năng lượng của tế bào và có thể gây ảnh hưởng đến các hoạt động sống khác của tế bào. Do đó, quá trình biểu hiện protein tái tổ hợp là hiệu quả hơn nếu quá trình đó có thể "bật" hoặc "tắt" theo yêu cầu. Vì lý do này, phần lớn các hệ thống biểu hiện được thiết kế theo kiểu được kích hoạt bởi việc bổ sung một phân tử nhỏ nào đó vào môi trường nuôi cấy vi khuẩn, thông thường là một cơ chất nào đó. Một số hệ thống biểu hiện khác được kích hoạt bởi nhiệt độ môi trường.

Các protein được biểu hiện trong tế bào vi khuẩn thường kém bền do sự biến tính gây ra bởi protease của tế bào chủ. Nhờ sự thay đổi trình tự của vector, ví dụ như việc bổ sung một hoặc một số amino acid vào đầu tận cùng N của phân tử protein, mà các protein tái tổ hợp có thể được bảo vệ tránh khỏi quá trình biến tính này. Các amino acid đó lại có thể được sử dụng như dấu hiệu để dễ dàng tinh sạch protein tái tổ hợp thông qua việc sử dụng các kháng thể tương ứng kết hợp với protein biểu hiện trên gel điện di. Các vector được thiết kế sao cho phần amino acid gắn thêm này có thể được loại bỏ nhờ phản ứng enzym hoặc phản ứng hóa học khỏi phân tử protein tái tổ hợp sau khi quá trình tinh sạch kết thúc.

Hiệu suất tổng hợp protein thường được tăng cao nhờ việc sử dụng các vector plasmid, bởi chúng có thể đồng thời tồn tại nhiều bản sao trong tế bào chủ. Tuy vậy, nếu số bản sao quá lớn thì có thể gây hiện tượng mất ổn định và plasmid có thể bị loại thải khỏi tế bào chủ: các tế bào này có xu hướng phát triển nhanh hơn các tế bào còn chứa plasmid và tỏ ra có ưu thế trong môi trường dẫn đến hiệu suất biểu hiện protein tổng cộng giảm đi. Để khắc phục tính mất ổn định của plasmid, người ta thiết kế một số vector có khả năng liên kết với nhiễm sắc thể của tế bào chủ, lúc đó chúng trở nên ổn định hơn, mặc dù hiệu suất biểu hiện của protein có thể bị giảm đi do số bản sao của gen tái tổ hợp ít hơn.

Hệ thống biểu hiện ở sinh vật nhân chuẩn

Mặc dù các hệ thống tế bào vi khuẩn có thể biểu hiện thành công nhiều loại protein có nguồn gốc sinh vật nhân chuẩn, nhưng có những loại protein thuộc kiểu này tỏ ra kém bền hoặc mất hoạt tính sinh học do hệ thống tế bào vi khuẩn thiếu ổn định hơn hoặc thiếu quá trình hoàn thiện protein sau

phiên mã, chẳng hạn như quá trình glycosyl hóa, giống như ở tế bào sinh vật nhân chuẩn. Hơn nữa, các loại protein được biểu hiện trong tế bào vi khuẩn đôi khi bị nhiễm với các hợp chất độc làm chúng trở nên không còn phù hợp để sử dụng làm thuốc cho người và động vật. Vì vậy, người ta đã xây dựng các hệ thống biểu hiện dựa trên các tế bào chủ có nguồn gốc từ sinh vật nhân chuẩn để có thể sản sinh ra các phân tử protein mang các đặc tính giống với dạng nguyên thủy của nó trong tự nhiên.

Các vector biểu hiện của sinh vật nhân chuẩn về cơ bản có nhiều đặc tính tương đồng với các vector có nguồn gốc vi khuẩn (*Hình 6.15*). Chúng cũng bao gồm một trình tự promotor, các trình tự tín hiệu điều khiển quá trình phiên mã và dịch mã. Một số thuộc tính bổ sung đặc trưng của vector sinh vật nhân chuẩn bao gồm các trình tự tín hiệu điều khiển quá trình hoàn thiện ARN và protein, chẳng hạn như quá trình bổ sung đuôi poly(A) sau phiên mã.

Các vector có thể tồn tại một cách độc lập với hệ gen tế bào chủ như trường hợp plasmid, hoặc liên kết vào nhiễm sắc thể của tế bào chủ. Một số vector được gọi là các vector con thoi có các trình tự của vi khuẩn cho phép việc thao tác và nhân vector diễn ra dễ dàng trong tế bào vi khuẩn, sau đó chúng được chuyển vào các tế bào sinh vật nhân chuẩn để biểu hiện protein.

Có ba loại tế bào chủ sinh vật nhân chuẩn được sử dụng: các tế bào nấm men, các tế bào côn trùng và tế bào động vật có vú.

Nấm men:

Nấm men *Saccharomyces cerevisiae* được sử dụng rộng rãi để biểu hiện các protein có nguồn gốc sinh vật nhân chuẩn. Một số đặc tính của nấm men này đặc biệt phù hợp với vai trò là tế bào chủ. Điều kiện nuôi cấy tế bào nấm men tương đối đơn giản và nhiều đặc điểm sinh lý và di truyền của chúng đã được nghiên cứu kỹ. Chúng có khả năng tiến hành quá trình hoàn thiện protein sau phiên mã và giải phóng protein ra môi trường giúp quá trình tinh sạch dễ dàng hơn. Hơn nữa, bản thân nấm men có cấu trúc plasmid gọi là 2μ plasmid có thể sử dụng làm vector tách dòng. ADN ngoại lai có thể biến nạp vào nấm men theo các phương pháp khác nhau, chẳng hạn thông qua quá trình dung hợp tế bào trần, hoặc nhờ kỹ thuật xung điện xử lý các tế bào dưới dòng điện ngắn có hiệu điện thế cao. Các vector biểu hiện của nấm men có thể tồn tại trong các tế bào như các plasmid (các vector episom) hoặc kết hợp vào hệ gen của tế bào chủ. Các vector episom có hiệu suất biểu hiện protein cao nhưng có xu hướng kém ổn định khi nuôi cấy ở quy mô lớn. Ví dụ về các protein tái tổ hợp được sản xuất bởi hệ thống tế bào nấm men là insulin của người được sử dụng trong điều trị bệnh nhân tiểu đường, và kháng nguyên bề mặt virus viêm gan B được dùng làm vaccin.

Các tế bào côn trùng:

Baculovirus là nhóm virus lây nhiễm ở động vật không xương sống, trong đó có các loài côn trùng. Các virus này thích nghi với tập tính sống sử dụng các vector biểu hiện của sinh vật nhân chuẩn. Các tế bào côn trùng được sử dụng làm tế bào chủ và đặc biệt thích hợp nhờ các protein do chúng tạo ra được glycosyl hóa giống với dạng tự nhiên của các protein động vật có vú. Baculovirus chứa một gen mã hóa cho một loại protein gọi là **polyhedrin** vốn được phiên mã dưới sự điều khiển của một promotor hoạt động cực mạnh. Do gen này không bắt buộc cần có cho sự sao chép hệ gen virus nên nó có thể được cắt bỏ và thay thế vào đó là gen tái tổ hợp được quan tâm. Vector của baculovirus chứa gen tái tổ hợp được chuyển vào tế bào chủ để

biểu hiện thành sản phẩm protein thông qua việc lây nhiễm. Các protein tái tổ hợp đã được sản xuất nhờ việc sử dụng hệ thống baculovirus có thể kể đến bao gồm hormon erythropoietin được sử dụng trong điều trị bệnh thiếu máu, hay hợp chất kháng virus β-interferon.

Các tế bào động vật có vú:

Các protein được sử dụng làm thuốc để điều trị bệnh ở người phải giống với dạng tự nhiên của nó. Trong nhiều trường hợp, protein tái tổ hợp chỉ được tổng hợp và hoàn thiện sau phiên mã một cách chính xác như dạng tự nhiên khi được biểu hiện trong tế bào chủ là tế bào động vật có vú. Vì lý do đó, người ta đã phát triển các hệ thống biểu hiện ở tế bào động vật có vú khác nhau. Các vector động vật có vú thường là các vector con thoi và có thể nhân lên dễ dàng trong tế bào vi khuẩn. Việc chuyển vector vào tế bào chủ thu được bằng kỹ thuật kết tủa Ca^{2+} hoặc bằng phương pháp xung điện. Nhiều loại protein tái tổ hợp khác nhau của người được sử dụng làm thuốc đã được sản xuất nhờ hệ thống biểu hiện của tế bào động vật có vú. Trong số các sản phẩm đó, có các hormon sinh trưởng, protein gây đông máu, yếu tố VIII, β-interferon và erythropoietin.

Kỹ thuật di truyền thực vật

Các phương pháp chọn, tạo giống truyền thống ở thực vật đã đóng góp nhiều thành công trong việc tạo ra nhiều giống cây trồng có năng suất cao, phẩm chất tốt. Tuy vậy, kỹ thuật di truyền hứa hẹn là phương pháp có thể trực tiếp cải biến các đặc tính sinh học. Các loài thực vật đặc biệt thích hợp cho việc tiến hành thao tác di truyền do các tế bào thực vật có **tính toàn năng**. Điều này có nghĩa là một cá thể thực vật hoàn chỉnh có thể được tái sinh từ một tế bào được cải biến di truyền duy nhất. Nếu cây tái sinh hữu thụ, thì gen mang đặc tính được cải tiến đó sẽ có trong hạt và vì vậy các cây mang đặc tính di truyền được cải tiến sẽ tiếp tục được sản sinh qua các thế hệ thông qua quá trình lai tạo.

Người ta đã phát triển các phương pháp khác nhau để tiến hành biến nạp gen ở thực vật. Trong đó, phương pháp điển hình được xây dựng dựa trên Ti plasmid vốn có sẵn trong vi khuẩn đất *Agrobacterium tumefaciens*. Trong tự nhiên, vi khuẩn này có khả năng gây nhiễm các loài cây hai lá mầm (nghĩa là tất cả các loài cây trồng trừ ngũ cốc) và gây nên sự phát triển của tế bào kiểu khối u, gọi là **nốt sần** ở rễ. Quá trình biến nạp gen vào tế bào thực vật là nhờ đặc tính của Ti plasmid. Trong Ti plasmid, có một vùng trình tự gọi là T-ADN có khả năng liên kết vào nhiễm sắc thể tế bào thực vật chủ. Sau đó, sự biểu hiện của các gen nằm trong vùng T-ADN được thực hiện trong tế bào chủ. Trên cơ sở nguyên lý đó, các vector chuyển gen ở thực vật được tạo ra từ Ti plasmid của vi khuẩn *A. tumefaciens*. Theo đó các gen ngoại lai được chuyển vào tế bào thực vật thông qua việc gắn những gen này vào vùng T-ADN của Ti plasmid. Hệ thống truyền gen này rất hiệu quả nhưng có hạn chế là chủ yếu được áp dụng cho các cây hai lá mầm vốn dễ dàng bị lây nhiễm bởi vi khuẩn. Ngoài ra, còn có một số phương pháp chuyển gen khác ở thực vật đã và đang được nghiên cứu và phát triển. Một trong những phương pháp đơn giản và cũng khá hiệu quả là phương pháp dùng "**súng bắn gen**". Trong đó, ADN thường được phủ trên bề mặt các hạt kim loại nặng có tính trơ cao như vàng và volfram, rồi được "bắn" vào tế bào thực vật nhờ áp lực của khí Heli. Những tế bào mang gen tái tổ hợp thành công sau đó sẽ được chọn lọc, tái sinh và sử dụng tiếp tục cho quá trình lai tạo và sản xuất giống.

Sau đây là một số đặc tính sinh học có giá trị đã được tạo ra nhờ kỹ thuật chuyển gen ở thực vật.

Tính kháng sâu, bệnh:

Nhờ kỹ thuật di truyền, các giống cây trồng khác nhau có khả năng kháng sâu, bệnh đã được tạo ra. Về cơ bản, có thể chia làm hai hướng chính. Hướng thứ nhất sử dụng việc chuyển một gen có nguồn gốc từ vi khuẩn *Bacillus thuringiensis* mã hóa cho một protein gọi là **protoxin** (hay còn gọi tắt là Bt) vốn gây độc ở nhiều loài sâu, bệnh. Gen này đã được biến nạp thành công vào cây cà chua, nhưng khả năng kháng bệnh của cây hạn chế do mức độ biểu hiện của gen Bt thấp. Hướng thứ hai liên quan đến việc làm thay đổi mức độ biểu hiện của các protein ức chế protease làm ảnh hưởng đến khả năng tiêu hóa mô thực vật của sâu, bệnh.

Tính kháng các bệnh virus:

Kỹ thuật di truyền cũng đã và đang được sử dụng để phát triển các cơ chế kháng bệnh virus mới. Chẳng hạn như, các trình tự gen đối mã tương ứng với gen virus được tiến hành biến nạp vào thực vật để ngăn cản sự biểu hiện hệ gen của virus. Tuy vậy hướng nghiên cứu này chỉ đạt được hiệu quả thành công trong một số trường hợp. Một phương pháp khác tỏ ra hiệu quả và thành công hơn liên quan đến việc chuyển một gen mã hóa cho protein vỏ virus. Mặc dù, đến nay cơ chế của tính kháng virus ở thực vật khi có sự biểu hiện của protein vỏ virus chưa được khẳng định chắc chắn.

Tính kháng thuốc diệt cỏ:

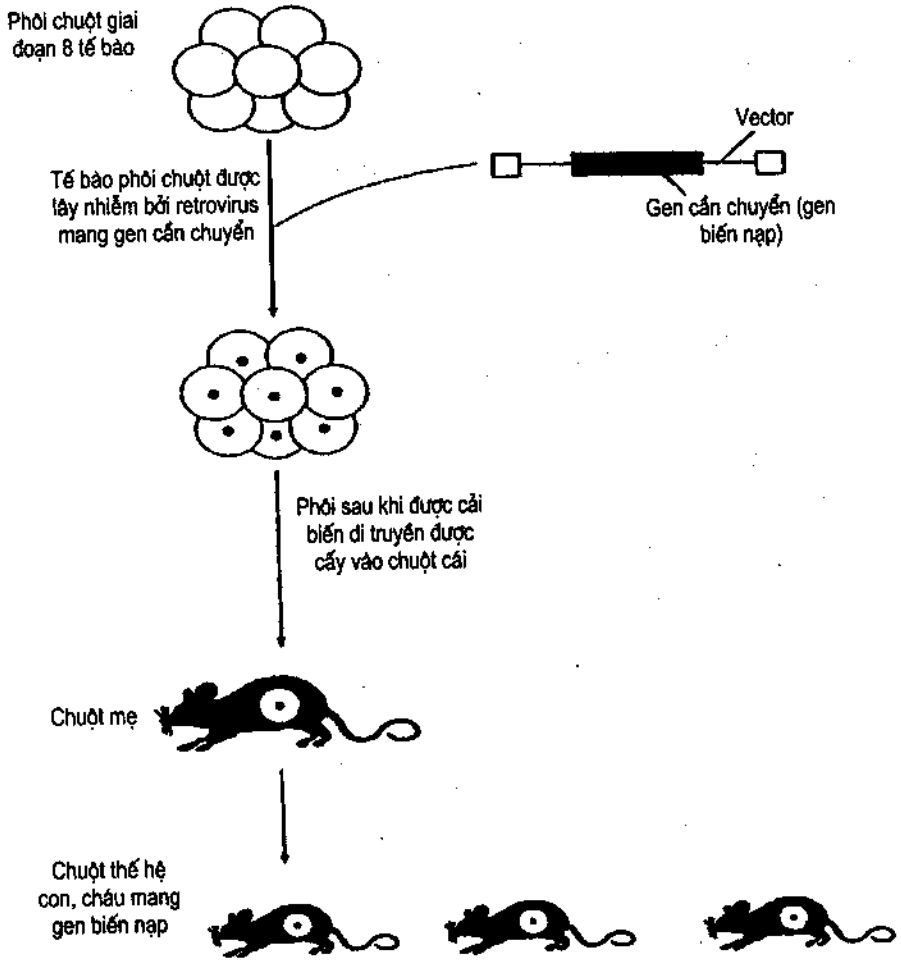
Mặc dù thuốc diệt cỏ được sử dụng rộng rãi, ước tính hàng năm khoảng 10% sản lượng nông nghiệp toàn cầu bị mất đi do sự phát triển của cỏ dại. Hơn nữa, phần lớn các thuốc diệt cỏ có giá thành đắt, độc với môi trường, đôi khi gây chết cây trồng và các loài cỏ khác. Kỹ thuật di truyền cho phép chuyển các gen kháng chất diệt cỏ vào thực vật, nhờ vậy, người ta có thể dễ dàng diệt các dạng cỏ dại một cách chọn lọc. **Glyphosate** là một chất diệt cỏ được sử dụng rộng rãi nhờ cơ chế ức chế enzym tổng hợp amino acid ở thực vật. Khả năng kháng với glyphosate được tạo ra nhờ việc chuyển gen mã hóa cho enzym tổng hợp amino acid không bị ảnh hưởng của chất diệt cỏ glyphosate.

Một số đặc tính nông học khác được cải tiến ở thực vật nhờ kỹ thuật di truyền gồm có khả năng chín chậm (ví dụ: ở cà chua) để thời gian bảo quản sau thu hoạch có thể kéo dài hơn, tính kháng các điều kiện bất lợi của môi trường (như hạn hán, đất nhiễm mặn, v.v...), hoặc làm tăng giá trị dinh dưỡng của sản phẩm. Do các tế bào thực vật thường dễ nuôi cấy và có khả năng tái sinh ở dạng sinh khối lớn, chúng có tiềm năng được sử dụng để sản xuất các loại protein tái tổ hợp. Vì vậy, một hướng nghiên cứu trong công nghệ sinh học là sử dụng các tế bào thực vật như các **nồi phản ứng sinh học** (bioreactor). Đã có những thành công nhất định theo hướng nghiên cứu này. Trong số đó, các kháng thể đơn dòng đã được sản xuất thành công từ các dòng tế bào thực vật được chuyển gen.

Động vật chuyển gen

Đến nay, các phương pháp chọn tạo giống động vật truyền thống đã được áp dụng rộng rãi để tạo ra các giống vật nuôi mang các đặc tính mong muốn như làm tăng năng suất sữa, tăng tốc độ sinh trưởng, v.v... Nhưng đối với nhiều tính trạng, việc lai tạo nhằm đạt được một tính trạng

mới thường làm thay đổi các tính trạng khác đang có. Hiện nay, với sự hỗ trợ của kỹ thuật di truyền, người ta có thể chuyển gen quy định một tính trạng mong muốn trực tiếp vào động vật. Các động vật được cải tiến di truyền theo phương pháp đó được gọi là **động vật chuyển gen** (transgenic animal), còn gen ngoại lai được gọi là **gen biến nạp** (transgene).



Hình 6.16. **Kỹ thuật chuyển gen ở động vật sử dụng vector retrovirus**

Các gen biến nạp có thể được chuyển vào tế bào động vật theo một số cách. Trong đó, các gen biến nạp được chuyển vào tế bào trứng đã thụ tinh hoặc vào tế bào ở giai đoạn đầu của quá trình phát triển của phôi. Các phôi sau khi được biến đổi được cấy truyền vào tử cung của một con mẹ, ở đó trứng mang gen tái tổ hợp sẽ phát triển thành một cá thể được cải biến di truyền.

Các vector retrovirus:

Các retrovirus có thể được sử dụng để tiến hành lây nhiễm các tế bào ở giai đoạn đầu của quá trình phát triển của phôi (*Hình 6.16*). Nhờ các vector retrovirus, các gen biến nạp được chuyển vào tế bào chủ và kết hợp với hệ gen nhân của tế bào chủ. Nhưng vì kích thước vector thường nhỏ, nên kích thước gen biến nạp thường bị hạn chế. Ngoài ra, việc sử dụng các vector retrovirus thường đòi hỏi các nguyên tắc an toàn sinh học.

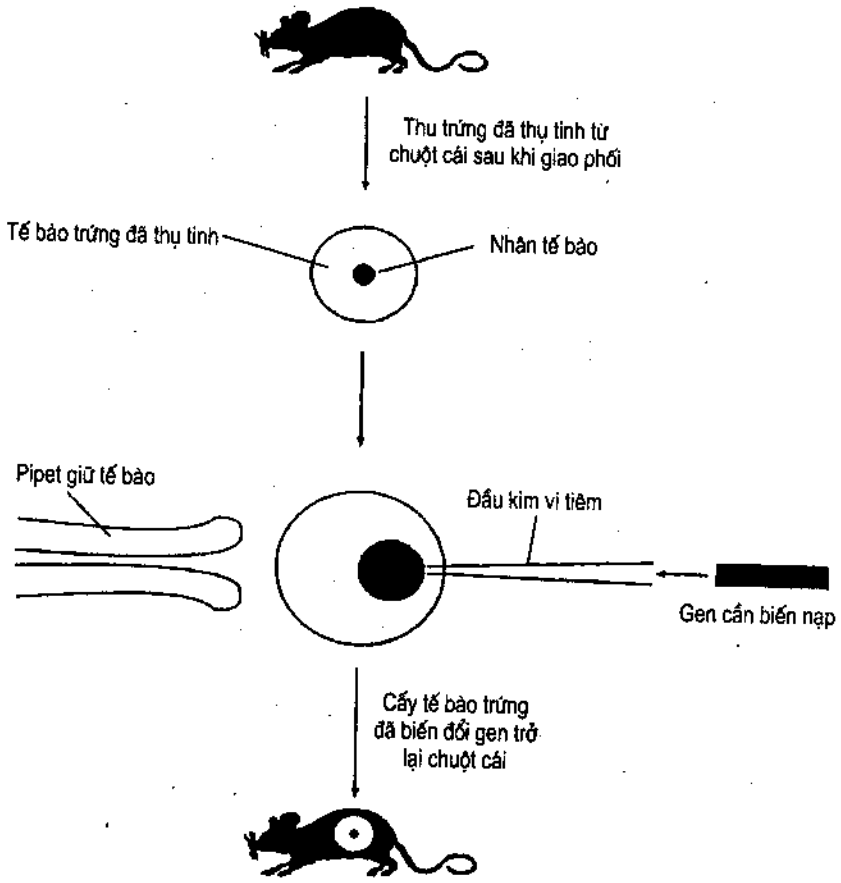
Phương pháp vi tiêm (microinjection):

Phương pháp này sử dụng việc tiêm trực tiếp ADN vào trong nhân của tế bào trứng đã thụ tinh dưới kính hiển vi (*Hình 6.17*). Mặc dù phương pháp này đòi hỏi sự tỷ mỷ, chính xác và có nhiều thao tác khó về mặt kỹ thuật, phương pháp này đến nay vẫn là phương pháp được sử dụng rộng rãi nhất ở động vật.

Các tế bào phôi gốc:

Các tế bào chuột ở giai đoạn đầu của quá trình phát triển của phôi có khả năng biệt hoá thành tất cả các loại tế bào khác được gọi là tế bào phôi gốc ES (*e*mbryogenic *s*tem cells). Những tế bào này có thể được tách ra và tiến hành nuôi cấy. Trong phòng thí nghiệm, người ta có thể biến nạp gen vào các tế bào ES rồi tiến hành cấy chuyển trở lại vào tử cung con cái đang mang thai để phôi tiếp tục phát triển thành cá thể động vật chuyển gen (*Hình 6.18*).

Các động vật chuyển gen được sử dụng làm công cụ trong các nghiên cứu và sản xuất protein tái tổ hợp. Ba ứng dụng chính của chúng có thể kể đến như sau:



Hình 6.17. **Chuyển gen ở động vật bằng kỹ thuật vi tiêm (microinjection)**

1) Tìm hiểu chức năng gen

Công nghệ chuyển gen ở chuột được hoàn thiện vào những năm 1980. Kể từ đó đến nay, người ta đã chuyển thành công hàng trăm gen vào chuột. Thông qua việc theo dõi các đặc tính của các cá thể biến nạp gen, người ta có thể tìm hiểu và phát hiện ra các chức năng của gen biến nạp. Điều này giúp làm sáng tỏ nhiều chức năng và cơ chế điều hòa hoạt động của gen.

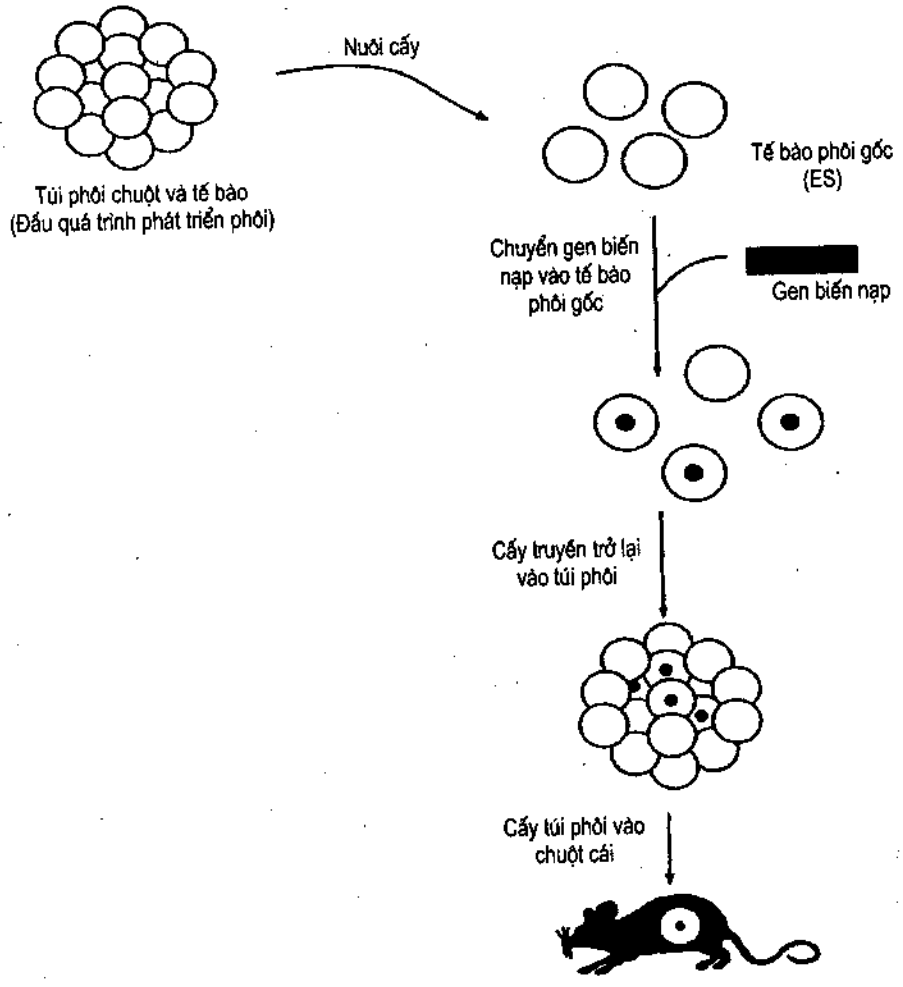
Một kỹ thuật liên quan đến định hướng nghiên cứu này là tiến hành chuyển gen để làm "bất hoạt" hoạt động của gen khác, gọi là kỹ thuật **knock-out**. Các cá thể có gen bị bất hoạt sẽ thiếu những chức năng nhất định, qua đó người ta có thể biết thông tin về chức năng của gen đó.

2) Mô hình nghiên cứu các bệnh ở người

Các động vật chuyển gen có thể được tạo ra để mô phỏng các bệnh ở người nhờ việc làm "hỏng" các gen nhất định nào đó. Toàn bộ mô hình nghiên cứu ở động vật có thể giúp nghiên cứu làm rõ cơ chế gây bệnh ở người, cũng như có thể sử dụng để sàng lọc và đánh giá các loại dược phẩm mới có tiềm năng chữa bệnh tương ứng. Hướng ứng dụng này đã từng được sử dụng để phát triển các mô hình nghiên cứu bệnh Alzheimer và chứng viêm khớp.

3) Sản xuất các protein tái tổ hợp

Kỹ thuật chuyển gen vào cừu và dê đã được thực hiện nhằm sản xuất một số loại protein người trong tuyến sữa. Việc sản xuất protein tái tổ hợp theo phương pháp này có một số ưu điểm. Sữa luôn được đổi mới, có thể được sản xuất với lượng lớn và việc thu sản phẩm không làm ảnh hưởng tới sức sống của con vật. Các loại protein được biểu hiện có đặc tính giống với dạng tự nhiên của người và dễ dàng tinh sạch do trong thành phần sữa chỉ chứa một số lượng ít các loại protein khác. Một số loại protein đã được sản xuất theo phương pháp này bao gồm protein yếu tố IX và α1-antitrypsin.



Hình 6.18. **Chuyển gen ở động vật thông qua tế bào phôi gốc (ES)**

TÀI LIỆU THAM KHẢO

1. A. G. Atherly, J. R. Girton, J. F. McDonal. *The Science of Genetics.* Saunders College Publishing (1999).
2. J. W. Dale, Malcolm von Schantz. *From Genes to Genomes: Concepts and Application of DNA Technology.* John Wiley & Sons Ltd (2002).
3. L. H. Danie, W. J. Elizabeth. *Genetics – Principles and Analysis.* Jones & Bartlett Publishers (1998).
4. A. J. F. Griffith, J. H. Miller, D. I. Suzuki, R. C. Lewontin, W. M. Gilbart. An *Introduction to Genetic analysis.* W.H. Feeman & Company (1996).
5. J. T. Hancock. *Molecular Genetics.* Butterworth-Heinemann Publisher (1999).
6. D. L. Hartl, E.W. Jones. *Genetics: Principles and Analysis (4th Ed.).* Jones & Bartlett Publishers (1998).
7. D. C. Hartl, D. Freifelder, L. A. Snyder. *Basic Genetics.* Jones & Bartlett Publisher (1988).
8. L. H. Hartwell et al. *Genetics: From Genes to Genomes.* McGraw-Hill (2000).
9. M. S. Kang. *Quantitative Genetics, Genomics and Plant Breeding.* CABI Publishing (2002).
10. H. M. Kingston. *ABC of Clinical Genetics (3rd Ed).*). BMJ Publishing Group (2002).
11. T. F. Kresina (Ed.). *An Introduction to Molecular Medicine and Gene Therapy.* Wiley - Liss, Inc. (2001).
12. H. H. Leland, Leroy Hood, L. G. Michael, L. R. Ann, M.S. Lee, C.V. Ruth. *Genetics: from genes to genomes* (2ndEd), McGraw-Hill (2004).
13. R. Lewis. *Human Henetics: Concepts and Applications.* McGraw-Hill (2003).
14. P. Paolella. *Introduction to Moleculr Biology.* McGraw-Hill (1998).
15. B. A. Pierce. *Genetics: A Conceptual Approach.* WH Freeman & Co. (2004).
16. R. J. Reece. *Analysis of Genes and Genomes.* John Wisley & Sons Ltd. (2004).

17. P. J. Russell. *Fundamentals of Genetics.* Addison Wesley Longman Inc, (2000).

18. D. P. Snustard, M.J. Simmons. *Principles of Genetics (3rd.).* John Wiley & Sons, Inc. (2003).

19. J. D. Watson et al. *Molecular Biology of the Gene (5thEd.).* Benjamin Cummings Press (2004).

20. R. F. Eeaver. P.W. Hedrick. *Genetics.* Wm. C. Brown Publishers (1997).

21. P. C. Winter, P.W. Hedrick, H.L. Fletcher. *Instant Notes in Genetics.* Bios Scientific Publishers (1998).

Chịu trách nhiệm xuất bản:
Chủ tịch HĐQT kiêm Tổng Giám đốc NGÔ TRẦN ÁI
Phó Tổng Giám đốc kiêm Tổng biên tập NGUYỄN QUÝ THAO

Tổ chức bản thảo và chịu trách nhiệm nội dung:
Chủ tịch HĐQT kiêm Giám đốc Công ty CP Sách ĐH – DN
TRẦN NHẬT TÂN

Biên tập và sửa bản in:
NGUYỄN HỒNG ÁNH

Trình bày bìa:
HOÀNG MẠNH DỨA

Chế bản:
QUANG CHÍNH

CHÚ GIẢI DI TRUYỀN HỌC
Mã số 7K701M7 - DAI
In 1.000 bản, khổ 19 x 27 cm, tại Công ty In - Thương mại TTXVN
70/342 Khương Đình - Hạ Đình - Thanh Xuân - Hà Nội.
Số xuất bản: 17 - 2007/CXB/ 91-2217/GD.
In xong và nộp lưu chiểu tháng 03 năm 2007.

CÔNG TY CỔ PHẦN SÁCH ĐẠI HỌC – DẠY NGHỀ
HEVOBCO
25 HÀN THUYÊN – HÀ NỘI
Website : www.hevobco.com.vn

TÌM ĐỌC SÁCH THAM KHẢO VỀ MÔI TRƯỜNG VÀ SINH HỌC CỦA NHÀ XUẤT BẢN GIÁO DỤC	
1. Chỉ thị sinh học môi trường	Bộ Giáo dục và Đào tạo
2. Kinh tế môi trường	PGS. TS. Hoàng Xuân Cơ
3. Cẩm nang Quản lý môi trường	PGS. TS. Lưu Đức Hải
4. Đất ngập nước	GS. TS. Lê Văn Khoa
5. Tài nguyên nước Việt Nam	PGS. TS. Nguyễn Thanh Sơn
6. Động vật có xương sống	GS. TS. Lê Vũ Khôi
7. Cơ sở hoá sinh	PGS. TS. Trịnh Lê Hùng
8. Chú giải di truyền học	Đỗ Lê Thăng, Đinh Đoàn Long

Bạn đọc có thể mua tại các Công ty Sách – Thiết bị trường học ở địa phương hoặc các Cửa hàng của Nhà xuất bản Giáo dục :

Tại Hà Nội : 25 Hàn Thuyên ; 187B Giảng Võ ; 23 Tràng Tiền ; 232 Tây Sơn.
Tại Đà Nẵng : 15 Nguyễn Chí Thanh ; 62 Nguyễn Chí Thanh.
Tại Thành phố Hồ Chí Minh : 240 Trần Bình Trọng – Quận 5 ;
104 Mai Thị Lựu – Quận 1 ;
Cửa hàng 451B, 453 Hai Bà Trưng – Quận 3.
Tại Thành phố Cần Thơ : Số 5/5, đường 30/4

8 934980 763384

Giá: 38.500 đ